Enactive / 07 .... Enaction_*in*_Arts

Proceedings of the
4th International Conference on Enactive Interfaces 2007

November 19 - 24, 2007

Grenoble, France

Organized by

Association pour la Création et la Recherche sur les Outils d'Expression
and
Laboratoire Informatique et Création Artistique

Chairs

Annie Luciani
Claude Cadoz

www.enactive2007.org

Enactive / 07
Proceedings of the 4th International Conference on Enactive Interfaces 2007

ISSN 1958-5497

Published and Edited by:
Association ACROE
INPG, 46 av. Félix Viallet
38000 Grenoble
France

Table of Contents

Foreword

Enactive / 07 is the Fourth International Conference on Enactive Interfaces, launched by the European Network of Excellence ENACTIVE.

In the continuation of previous editions (2004, Villard-de-Lans, France; 2005, Genoa, Italy; 2006, Montpellier, France), Enactive / 07 aims at promoting the concept of Enaction in the field of Information and Communication Technologies. Creative researchers, anticipative theorists, innovative engineers and producers are invited to confront their last theoretical, experimental, technological and applications advances during various talks, demos and posters sessions.

The aim of the conference is to encourage the emergence of a multidisciplinary research community in a new field of research and on a new generation of human-computer interfaces called Enactive Interfaces. Authors have proposed papers in all the branches covering Enactive Interfaces, Interaction and systems, including but not limited to:

- Technology of interaction, virtual realities, robotics, interactive computer graphics and sound synthesis, multimodal and gesture interaction, haptic systems;
- Psychology of perception and action, cognitive sciences, cognitive ergonomics;
- Philosophy and epistemology of Enaction;
- Applications to creativity, arts and design, learning and teaching, industrial design, special users and uses, sport and entertainment.

Seizing the opportunity of its organization in Grenoble, an historical place in France for innovation in Arts and Culture, the 4th International Conference on Enactive Interfaces is exceptionally extended by a scientific and artistic event **Enaction_*in*_Arts**. Arts and Culture are main human activities that are fed by the concepts of Enaction and of Enactive Knowledge. In Arts, concepts and models are increasingly evolving towards the rehabilitation of performance and interaction as principal components of contemporary artistic activity. Within Enactive / 07, Enaction_*in*_Arts proposes:

- a one-day symposium in which artists and researchers investigate the shifts that the concept of Enaction causes in Arts and Culture in the context of Information Society Technologies, at the philosophical, technological and artistic levels;
- innovative forms of Arts: multimedia concerts merging music, visual arts and interactive choreography, interactive artistic installations including virtual realities and robotics;
- a pioneering exhibition entitled “Touch the Future”, offering to the grand public interactive demos that are emblematic of Enaction and Enactive Interfaces.

These Proceedings present the state of the art in the theory, technologies and applications of Enaction, in a wide variety of topics and disciplines. They are organized in two parts:

- **Enactive / 07 Papers**: a part dedicated to the general scientific topics related to Enaction, in continuation of the past conferences;
- **Enaction_*in*_Arts Papers**: a part specifically devoted to Enaction_*in*_Arts-oriented papers.

Many thanks are due:

In the first place, we wish to thank all the scientists who contributed to this volume and participated to the conference, and we express our gratitude and encouragement to the authors whose submissions were not successful. We express our gratitude to our co-chairs for their help in preparing the scientific program of Enactive / 07 .... Enaction_*in*_Arts, and to the members of the review committee who have evaluated the contributions. We wish to thank the members of the interdisciplinary scientific and artistic committee of Enaction_*in*_Arts, who selected artworks in so

various formats. They are too numerous to appear here but their names are listed in this volume. To all, thank you for having contributed to this conference.

Support to the Fourth International Conference on Enactive Interfaces was provided by the European Network of Excellence ENACTIVE, by the *Ministère de la culture et de la communication*, the *Ministère de l'Education Natianale, de l'Enseignement Supérieur et de la Recherche*, the *Ministère de l'Industrie*, the *Conseil Régional Rhône-Alpes*, the *Conseil Général de l'Isère*, *Grenoble-Alpes Metropole*, the *Ville de Grenoble*, the *Institut National Polytechnique de Grenoble*, the *Université Joseph Fourier*. Without these sources of funding, neither the conference nor the publication of the present volume would have been possible.

The organization had benefited from the partnerships with *Canada Council of Arts*, *Canada Council of Letters, Statens Kunstfond of Danemark, SACEM*, and from the active support of local cultural organisations: *Maison de la Culture de Grenoble, Conservatoire de musique et de danse de Grenoble*, *Régie Grenoble Bastille*. Our industrials partners, *Acquisys*, *Immersion*, *Papillon*, *PianoTech*, *Concurrent Computer Coorporation*, by their help, contributed actively to the demos, exhibitions and artworks presentations.

Olivier Tache (web site, contact with authors, proceedings, design), Nicolas Castagné (Minatec hall, sponsors), Aurélie Arliaud (concerts), Kevin Sillam (interactive installations), Matthieu Evrard ("Touch the Future" exhibition), Olivier Kahn (communication and media), Ali Allaoui and Damien Couroussé (sponsors), and Maria Guglielmi (general secretary) have contributed tremendously to the preparation of this event. Without them, this conference would not have been possible.

We also express our gratitude to the other members of the ICA laboratory and of ACROE (Julien Castet, Geoffrey Cerna, Daniela Favaretto, Jean-Loup Florens, Sylvain Marlière, François Poyer) for their permanent help in organizing Enactive / 07 .... Enaction_*in*_Arts.

Annie Luciani,
Claude Cadoz,
Organizers of Enactive / 07 .... Enaction_*in*_Arts

Conference Committees

Chairs

Annie Luciani
Laboratoire ICA, Grenoble, France
Claude Cadoz
ACROE & Laboratoire ICA, Grenoble, France

Enactive/07 Co-Chairs

Philosophy

Charles Lenay
COSTECH, France
Elena Pasquinelli
Institut Nicod, France

Technology

Ronan Boulic
EPFL, Switzerland
Chris Raymaekers
UHasselt, Belgium

Human action and perception

Benoît Bardy
UM-1, France
Thomas Stoffregen
HFRL, USA

Applications and uses

Carlo Alberto Avizzano
PERCRO, Italy
Charlotte Magnusson
ULUND, Sweden

Hands-on demos and tutorials

Ian Summers
UNEXE, UK
Teresa Gutiérrez
LABEIN, Spain
Emilio Sanchez
CEIT, Spain

Enactive/07 Review Committee

Carlo Alberto Avizzano
Ali Allaoui
Benoît Bardy
Massimo Bergamasco
Joan De Boeck
Ronan Boulic
Antonio Camurri
Sylvain Cardin
Marcello Carrozzino
Nicolas Castagné
Julien Castet
Parag Chaudhuri
Damien Couroussé
Pierre Davy
Gunnar Declerck
Didier Delignières
Matthieu Evrard
Haakon Faste
Jean-Loup Florens
Olivier Gapenne

Amalia De Goetzen
Jean-Baptiste Guignard
Teresa Gutiérrez
Chimin Hsieh
Thomas Hulin
Gunnar Jansson
Klaus Joehl
Armen Khatchatourov
Alexandros Kontogeorgakopoulos
Charles Lenay
Annie Luciani
Charlotte Magnusson
Ludovic Marin
Sylvain Marlière
Luca Mion
Denis Mottet
Manfred Nusseck
Sile O'Modhrain
Renaud Ott
Elena Pasquinelli

Achille Peternier
Giovanni De Poli
François Poyer
Sofiane Ramdani
Chris Raymaekers
Emanuele Ruffaldi
Emilio Sanchez
Francois Sebbah
Pierre Steiner
John Stewart
Thomas Stoffregen
Ian Summers
Olivier Tache
Franco Tecchia
Hugues Vinet
Marcello Wanderley

Enaction_*in*_Arts Artistic and Scientific Committee

Hubertus von Ameluxen
EESI Angoulême-Poitiers, France
Philippe Baudelot
Digital Arts & Cultures, Nice, France
Massimo Bergamasco
PERCRO, Italy
Ludger Brümmer
ZKM, Germany
Elise Bureau
LIA, France
Antonio Camurri
DIST, Italy
Annie Luciani
INP-G, France
Sile O'Modhrain
QUB, Northern Ireland, UK
Arnaud Petit
CNR de Grenoble, France
Nicolas Tixier
Ecole nationale supérieure d'architecture de Grenoble, France
Hugues Vinet
IRCAM, France
Marcelo Wanderley
Mc Gill University, Canada

Organizing Committee

Ali Allaoui
Aurélie Arliaud
Claude Cadoz
Nicolas Castagné
Julien Castet
Geoffrey Cerna
Damien Couroussé
Matthieu Evrard
Daniela Favaretto
Jean-Loup Florens
Maria Guglielmi
Maxime Houot
Chimin Hsieh
Olivier Le Van
Annie Luciani
Sylvain Marlière
François Poyer
Kevin Sillam
Olivier Tache

Sponsors

The organization of ENACTIVE/07 .... Enaction_*in*_Arts was supported by:

ENACTIVE / 07 Papers

KEYNOTE

-

From Actual to Virtual Action

Jean-Pierre Roll

Université de Provence / CNRS – UMR 6149

Pôle 3C

Laboratoire de Neurobiologie Humaine - Centre St Charles

3 Place Victor Hugo - 13331 Marseille Cedex 03

E-mail: Jean-Pierre.Roll@up.univ-mrs.fr

Abstract

Patients with proprioceptive loss are totally deprived of the ability to perceive their own body, and in addition, they can no longer move. This pathological model definitely proves the existence of a "circular loop" between motor representation and motor control. Do we enact our "kinaesthetic experience", since body ownership is not consciously acquired at once but gradually set up as the result of active experience?

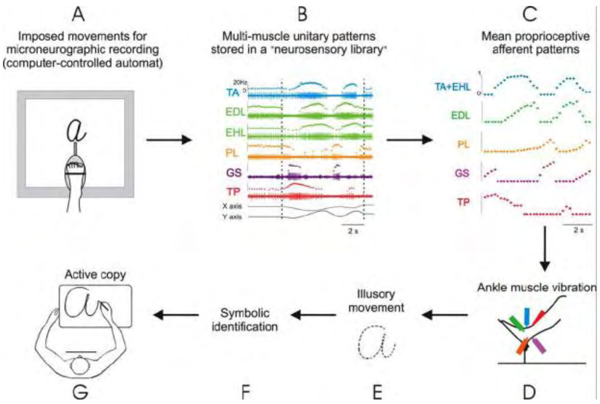
Proprioceptive cues contribute crucially to body awareness as well as providing a link between the body and extrapersonal space. The network responsible for mechanosensitivity is deeply buried and spreads throughout the whole musculature, and the complex structural and functional properties of the proprioceptive system have only recently been brought to light.

The proprioceptive sensory codes are oriented and weighted population codes, as recently established in humans using microneurographic methods.

The proprioceptive feedback generated by complex movements results from the exactly adjusted and timed lengthening and shortening of all the muscles crossing the joints during their execution. The afferent patterns thus produced are specific to each of our actions: they therefore constitute a set of "neurosensory bar-codes", each of which is the "sensory signature" of a given action. These patterns are highly specific and perfectly reproducible when the same action is repeated.

Since we have been collecting these proprioceptive afferent patterns for several years at our laboratory, we have built up a proper "human neurosensory library" containing the proprioceptive signatures of various movements with trajectories corresponding to written symbols. By applying to human subjects vibration patterns which were copies of "natural" afferent responses, it was found to be possible to evoke the illusion of the corresponding movements although no actual movements were performed. In addition, the fact that the subjects were able to recognize and name the symbols corresponding to the virtual movements evoked by the vibration suggests that "proprioceptive" and "kinaesthetic" signatures may evoke the "cognitive signature" associated with our actions, or at least with those involved in forming symbolic shapes.

We are now working on the development of a "proprioceptive generator of Virtual Movements" in humans: this should constitute a useful tool in the fields of motor learning, rehabilitation and virtual reality.



Mean stages in the experimental design.

A In previous studies, microneurographic muscle spindle afferent recordings were performed at the level of the lateral peroneal nerve during passive movements describing letters or numbers imposed on the ankle joint.

B Examples of unitary Ia afferent responses produced by each of the six main ankle muscle groups.

C The response of each population of muscle afferents was averaged for each graphic sign (here, the letter a).

D The averaged muscle responses were used as patterns to pilot the vibrators applied to each corresponding muscle tendon.
E The subjects perceived an illusory movement.
F They first had to recognize and name the graphic sign.
G They then had to draw the perceived trajectory on a digitizing tablet

References

[1] Bergenheim M, Ribot-Ciscar E, Roll JP (2000) Proprioceptive population coding of 2-D movements in humans: Part I. Muscle spindle feedback during "spatially oriented movements" Exp. Brain Res., 134: 301-310.

[2] Roll JP, Bergenheim M, Ribot-Ciscar E (2000) Proprioceptive population coding of 2-D limb movements in humans: Part II. Muscle spindle feedback during "drawing like movements" Experimental Brain Research, 134: 311-321.

[3] Ribot-Ciscar E, Bergenheim M, Roll JP (2002) The preferred sensory direction of muscle spindle primary endings influences the velocity coding of two-dimensional limb movements in humans. Exp. Brain Res., 145(4): pp 429-36.

[4] Ribot-Ciscar E, Bergenheim M, Albert F, Roll JP (2003) Proprioceptive population coding of limb position in humans. Exp Brain Res. 149: 512-519

[5] Roll JP, Albert F, Ribot-Ciscar E, Bergenheim M (2004) "Proprioceptive signature" of cursive writing in humans: a multi-population coding. Exp Brain Res. 157: 359-368

[6] Albert F, Ribot-Ciscar E, Fiocchi M, Bergenheim M, Roll JP (2005) Proprioceptive feedback in humans espresses invariant during writing. Exp Brain Res. 164: 242-249

[7] Albert F., Bergenheim M., Ribot-Ciscar E., Roll J.P. (2006) The Ia afferent feedback of a given movement evokes the illusion of the same movement when returned to the subject via muscle tendon vibration. Exp. Brain Res., 19: 1-12

[8] Roll JP (2003) Physiologie de la Kinesthèse "La proprioception musculaire: Sixième Sens ou Sens Premier ?" Intellectica. Vol. 36/37 - pp 49-66.

KEYNOTE
-
All About Thresholds: An Overview of Human Haptic Perception of Mechanical Properties

Hong Z. Tan
Haptic Interface Research Laboratory
Purdue University
465 Northwestern Avenue
West Lafayette, IN 47907-2035
E-mail: hongtan@purdue.edu

Abstract

Enactive interfaces capitalize on human's exquisite sensitivity to object properties through physical interactions with the environment. While such interactions are natural and intuitive, our sensitivity to mechanical properties needs to be specified explicitly so that the enactive interfaces can successfully match the human sensory-motor capabilities. In this talk, I will summarize what we currently know about the human discrimination thresholds for displacement, joint-angle position, force magnitude and direction, and stiffness. These data can be quite useful for solving engineering problems, as I will illustrate with examples. I will end the talk with our recent results on human haptic perception of complex waveforms using, again, threshold measurements. I hope that the threshold data summarized in this overview will be informative to the designers of enactive interfaces.

Biography

Hong Z. Tan is currently an associate professor of electrical and computer engineering at Purdue University in West Lafayette, Indiana. She received her Bachelor's degree in Biomedical Engineering from Shanghai Jiao Tong University, P.R. China. She earned her Master and Doctorate degrees, both in Electrical Engineering and Computer Science, from the Massachusetts Institute of Technology (MIT). She was a Research Scientist at the MIT Media Laboratory before joining the faculty at Purdue's School of Electrical and Computer Engineering in 1998. She currently holds a courtesy appointment in the School of Mechanical Engineering and the Department of Psychological Sciences at Purdue University. She is also a Faculty Fellow in the Envision Center for Data Perceptualization at Purdue University. Tan has held a McDonnell Visiting Fellowship at Oxford University, and a Visiting Associate Professorship in the Department of Computer Science at Stanford University.

Tan's research interest is in the area of haptic human-machine interface and human haptic perception. She has published more than 100 research articles in journals, conference proceedings and books. She was a recipient of the US National Science Foundation's Early Faculty Development (CAREER) Award from 2000 to 2004. In addition to serving on numerous conference program committees, she was a co-organizer (with Blake Hannaford) of the International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems from 2003 to 2005. She currently serves as the Chair of the IEEE Technical Committee on Haptics, a home for the international interdisciplinary haptics research community.

Enactive interface in Simulation of Medical Manipulation ICSI

Norihiro Abe* Ryutaro Mizokami* Yoshimasa Kinosita*‡ Shoujie He‡
(*)Kyushu Institute of Technology, JAPAN
(†), Munakata Suikokai General Hospital, JAPAN*
(‡)Eastman Kodak, USA
E-mail:abe@mse.kyutech.ac.jp, shoujie.he@gmail.com, yosimasa@suikokai.or.jp

Abstract

A remote surgery approach is not yet practical, because it solely relies on the visual perception with no input for the haptic perception. Since it is a lot faster for a human being to sense through touching than by watching especially when he is manipulating an object, sensing through touching is considered essential to the success of remote surgery. The system proposed simulates the process of ICSI. A force sensing system PHANToM is employed for collecting the information for the haptic perception. It provides an operator with force input in 6 DOF and force output in 3 DOF. OpenGL is used for the 3D graphics rendering. By synchronizing the information for both the visual and haptic perception, a virtual reality system that reacts to both the visual and force interactions has been developed. The simulation results reported in this paper proved that using the force sensing system makes a significant difference in the medical treatment.

1. Introduction

Recently, micro cells are becoming the point of focus in the research of medical treatment technologies. Intracytoplasmic sperm injection (ICSI) is one of the technologies. The ICSI is an in vitro fertilization procedure, which is most commonly used to overcome male infertility problems. The procedure is done under a microscope using micromanipulation devices. A hollow needle loaded with a single sperm is first injected into an ovum. The sperm will then be released into the cytoplasm of the ovum. One of the problems with the ICSI technology is that the cell structure of the ovum could easily be destructed. Once the cell structure of the ovum is destructed, all kinds of issues such as low quality of the divided ovum, failure of imbedding, or miscarriage even after imbedding could occur. Therefore, the ICSI technicians are required to be skillful enough not to destruct the cell structure of the ovum during the ICSI process. Another technology that is becoming popular over the past few years is the technology for the remote surgery with the help of a robot-based surgery system. With this technology, it is possible for a patient with heart disease and hospitalized in a regional hospital to receive the treatment from a specialist in a large hospital of a large city. It is expected that the regional difference in terms of the quality of healthcare in the area of surgery could be reduced through the remote operations with the help of such a robot-based surgery system. The basic idea of this technology is to transmit the images of the patient in a fast speed so that the surgeon could do the surgery remotely by examining the images.

The above-mentioned two technologies, however, require the operator to manipulate the target objects with no input for the haptic perception. On the other hand, surgeons and ICSI specialists in general heavily rely on the haptic perception while identifying organ and tools for a surgery. Diagnosing by touching is also a common practice in medical treatment. In another words, without the haptic perception, surgeons and the ICSI specialists are required to be highly skillful by working with their visual perception only.

It is a lot faster for a human being to sense through touching than by watching especially when he is manipulating an object. Visual perception on average takes 30msec to finish processing the incoming visual information. Haptic perception, on the other hand, takes only 1msec to finish processing the information sensed by touching. It is foreseeable that the failures and inappropriate operations that occurred when only visual perception was available could be largely avoided by incorporating the information processed by the haptic perception.

This research aims at proving the significance of employing force sensor in medical treatment. As the very first step, a simulation system is developed, which simulates the ICSI in a virtual space with the force sensor added. Experimental results show that using the information collected through the force sensor makes a significant difference in terms of the quality of treatment.

2. System Organization

The simulation system consists of three subsystems, interface, force computing and rendering subsystem. The interface subsystem collects the data through the force sensor while user manipulates the needle of the

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

ICSI. The force computing subsystem evaluates the data collected by the interface subsystem and determines the force information to be passed along to the user in the virtual space. The rendering subsystem visualizes the entire simulation process to the user. In this system, PHANToM premium 1.0 manufactured by SensAble Technologies Inc. is used in the interface subsystem for collecting the force data. Although only one PC is used, force computing and rendering subsystem work concurrently.

3. Information Synchronization

3.1. Process flow

The rendering and force computing happen concurrently. The former updates the information in a frequency of 30Hz, while the latter does in 1kHz. Synchronized with these frequencies, the visual information and the force information are perceived by human being in real time.

3.2. Modeling of the haptic perception

Haptic perception is an important means of feedback. The skin of human body contains three types of sensory receptors: thermoreceptors responding to heat and cold, noiceptors responding to which intense pressure, heat and pain, and mechanoreceptors responding to pressure. The last type is of interest to this research.

When a human operator is doing a manual operation, the haptic perception of his hands keeps receiving the stimuli. Throughout the operation, he adjusts his movement or power according to the feedback from his haptic perception. First of all, he needs to sense the shape and hardness of an object at the time of grasping the object. He may then need to decrease or increase the pressure applied to the object if the object is soft. In the medical treatment, many operations reply on the haptic perception as well. For example, checking the progress of an operation needs the haptic perception. Evidently, information collected through the haptic perception is useful and significant to many real-world operations. Therefore, when we simulate the real-world operations in the virtual space, information for the haptic perception should also be incorporated in addition to the information for the visual perception.

Here are some details about how to model the force sensing process. The ovum simulated in this system is assumed to be a visco-elastic object. Under this assumption, the relationship between the force bounced by the ovum F and the deformation quantity Δx could be represented in equation (1).

$$F = -(c\Delta\dot{x} + k\Delta x) \qquad (1)$$

Fig. 1 illustrates the deformation quantity of the ovum. Fig. 2 illustrates the visco-elastic model. The strength of force F is measured by PHANToM. Once the membrane of the ovum is penetrated, will there be a frictional force between the surface of the needle model and the membrane of the ovum model. The frictional force is also simulated and its strength is measured by the PHANToM as well.

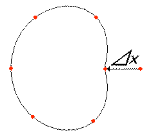


Fig. 1 Deformation of the ovum

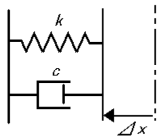


Fig. 2. The visco-elastic model.

3.3. Modeling of the visual perception

OpenGL is used for rendering the objects simulated in this system. The primitive geometric models provided by the OpenGL and their combination are used for the modeling of the needle and the pipette used for stabilizing the ovum. Non-uniform Rational B-Splines (NURBS) curved surface is used for the modeling of the ovum. With the NURBS curved surface, it is possible to render the deformed model by moving the control point. Modeling an ovum with the NURBS curved surfaces has the following steps involved.

1. Create a ball.
2. Divide the ball in a radial pattern
3. Replace each of the patches with the NURBS curved surface.
4. Merging the models

GHOST is employed for the synchronization between OpenGL based modeling and PHANToM based measuring of the force data. As a result, the information for visual perception and the information for haptic perception are successfully perceived synchronized. Merging all the models described above in a well-synchronized fashion, it is possible to have both the visual perception and the haptic perception enabled in simulating the deformation process of an ovum in a virtual space.

4. Simulation System

Fig. 3 shows a simulation window, which is a snapshot of the virtual space created by the simulation system. The ball shaped graphics in the center of the window is the model of an ovum. This model is rendered with 20 NURBS curved surfaces. The stick shaped graphics on the right side is the model of a needle. This needle model could be moved back and forth by an operator working on the PHANToM. Any force resistance from the virtual space will be sent back to the user also through the PHANToM. The cylinder shaped graphics on the left side of the window is the model of a pipette used for stabilizing the ovum. The graphics for the pipette model is rendered with a combination of the primitive functions provided by OpenGL.

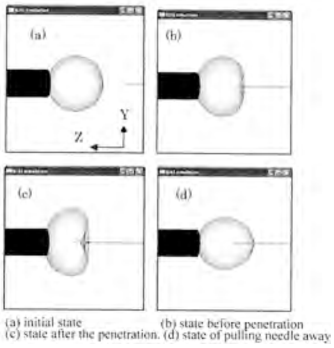


Fig. 3. Process of ICSI simulation.

5. Experimental Results

5.1. Experiment 1

5.1.1. Methodology

With the simulation system, the following sequence of actions was taken. During each action, the deformed amount of the ovum is measured and the force resistance along the Z-axis is also measured through the PHANToM.

Action 1: Move the needle in a speed of 0.4µm/msec towards the positive side of the Z-axis until the membrane of the ovum is penetrated.

Action 2: Once the membrane of the ovum is penetrated, immediately stop moving the needle until the penetrated membrane stopped its movement.

Action 3: As soon as the penetrated membrane stopped its movement, start moving the needle in a speed of 0.4µm/msec towards the negative side of the Z-axis.

5.2. Experimental Results

The graphics corresponding to the sequence of the three actions mentioned in Section A-1 is given in Fig. 4. Fig. 5 shows a graph of the force data measured by the PHANToM.

Fig. 4(A) shows the state right before the simulation starts. Fig. 4(B) shows a snapshot of Action 1. Fig. 4(C) shows a snapshot of Action 2. Fig. 4(D) shows a snapshot of Action 3.

In the graph shown in Fig. 5, the time periods of (a) and (b) covered the entire process of Action 1. The time period (c) is when Action 2 is being taken. The time periods (d), (e), and (f) covered the entire process of Action 3.

The time periods of (a) through (f) correspond to the following states.

Period (a): the state before the needle reaches the ovum yet. There is no force added to the needle.

Period (b): the state while the needle is pushing against the ovum but the membrane of the ovum has not been penetrated yet. There is a force resistance from the ovum.

Period (c): the state right after the membrane of the ovum is penetrated. There is a frictional force while penetrated membrane is moving along the needle.

Period (d): the state while the needle is being moved away from the ovum after the movement of membrane stops. There is a force resistance from the membrane. Therefore, there is no relative movement between the needle and the membrane during this time period.

Period (e): the state when the needle starts moving out of the ovum. Since there is relative movement between the needle and the membrane, there is a frictional force between the membrane and the needle.

Period (f): the state after the needle has been fully pulled out from the ovum. There is no force added to the needle.

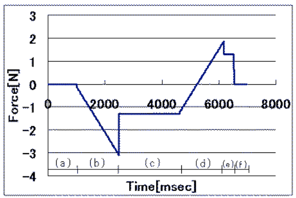


Fig.4. Needle insertion modeling

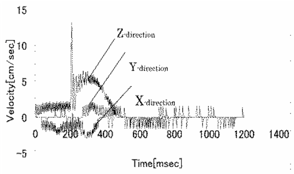


Fig. 5.Velocity in setting 1.

5.3. Experiment 2

5.3.1. Methodology

The experiment of using PHANToM to conduct the ICSI procedure is done with the following three settings. The speed data of PHANToM are collected throughout the ICSI process.

Setting 1: Information for both visual perception and the haptic perception is made available throughout the entire ICSI process.

Setting2: Information for both visual perception and haptic perception is made available until the needle reached the ovum. After the needle reached the ovum, only the information for the haptic perception is made available.

Setting3: Only the information for visual perception is made available throughout the entire ICSI process.

5.3.2. Experimental Results

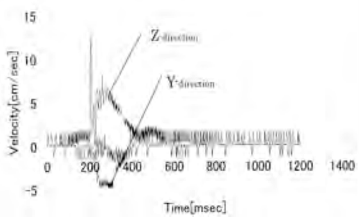
Fig. 5 to 7 show the speed data collected throughout the ICSI process for the three settings. The penetration timing is at 200ms.

6. Discussion

With the simulation results in Fig. 3, it can be concluded that the entire deformation process of the ovum model has been successfully simulated in Experiment 1.

The detailed analysis about the speed changes of PHANToM shown in Fig. 5 to 7 will be given in the following context. Before that, it is necessary to introduce a concept of delay time, which is a period of time between the timing when the penetration happens and the timing when the movement of PHANToM stops.

At the time when t ms elapsed since the penetration happened, if the average speed between t and t+0.01 sec is less than 0.001 cm/sec, it is considered that the movement of the needle has stopped. The time period of t is referred to as the delay time.



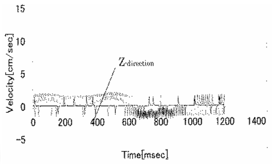


Fig.6.Velocity in setting 2   Fig. 7. Velocity in setting 3.

It can be concluded that when it is setting 1, the simulation responded fastest. It can also be seen that with either the information for visual perception only or the information for haptic perception only, the state of the ovum can still be perceived although the delay time is relatively long. Eventually, the movement of PHANToM stopped.

From the graphs in Fig. 5 and 6, it can be seen that when it is with setting 1 and 2, PHANToM moved in a sharply increased speed along the Z-axis immediately after the penetration happened. This is because the force resistance from the membrane of the ovum reduced substantially at the time that the membrane was penetrated. On the other hand, when it is with setting 3, since there is no haptic perception throughout the ICSI process, the speed along Z-axis constantly stayed at 1 cm/sec. Even after the penetration happened, there was no significant change in the speed and the PHANToM stopped with a delay time of 638 ms. In setting 1, the delay time was 516 ms. Taking into account the substantially increased speed right after the penetration, it can be concluded that with the haptic perception, the delay time of 638 ms in setting 3 could be shortened to 516 ms, which is the shortest possible delay time.

The experimental results from Experiment 2 indicate that the ovum state could be perceived even if only the haptic perception is used. Comparing the graphs of setting 1 and 3 (Fig. 5 and 7), however, collecting force data for the haptic perception is actually functioning against the ICSI operation in this simulation. Although it is important to sense the state of an object with visual perception and haptic perception, inappropriate force data collection could play an opposite role. The problem is that the force changes before and after penetration was set too large. Therefore, an optimal way of force data collection should be investigates so that the haptic perception could be best used.

Experimental results with setting 2 proved that it is possible to perceive the state of the ovum even with the haptic perception only. When the information for visual perception is not available in a treatment, the haptic perception could be considered as an alternative.

7. Conclusion

The efficiency and effectiveness of synchronizing the information collected from the force sensor with the visual information have been proved. Force data collection for haptic perception is very important for sensing the state of an environment. It could play an opposite role as well if an appropriate way of force data collection is not in place. It is necessary to investigate when and where the force data collection for the haptic perception should be introduced. When the information for visual perception is not sufficient, it will be the perfect time to add the haptic perception. For instance, in the surgery of an aneurysm in brain, using haptic perception will make significant difference, because it normally locates deep in the brain and it is impossible to remove other tissue or organ to make the aneurysm visible. A surgery simulation system with insufficient visual perception will be researched and developed. In order to do so, it is necessary to reconstruct the organ from CT images and come up with real-time algorithms for simulating the process of organ deformation.

Acknowledgement

We greatly appreciate the aid of Ministry of Internal Affairs and Communications (MIC) and the Grant-in-Aid for Scientific Research (A) and (S).

References

[1] S. Cotin, H. Delinguette, and N. Ayache. Real-time elastic deformations of soft tissues for surgery simulation, In IEEE Transactions On Visualization and Computer Graphics, 5(1), pages 62-73, January-March 1999

[2] Yasuhiro Watanabe, Daisuke Tokumochi, Norihiro Abe, Kazuaki Tanaka, Hirokazu Taki, Yoshimasa Kinoshita: Cutting Virtual Organ Model with Haptic Feedback Device, -CME2005, pp.255-260

[3] Koich Yamamoto, Kazuaki Tanaka, Norihiro Abe, Yoshimasa Kinosita, Akira Yokota : Cutting operation of virtual object and it application to medical simulation, ICAT'99, pp.161-165, 1999

Enactive Art:
Parietal *and* Frontal Brain Art?
From Pictorial to Speech Evidence

Christian ABRY, Marie-Agnès CATHIARD, Marie-Laure DIAFERIA*

Center for Research on the Imaginery (CRI), Université Stendhal, Grenoble (France),

**also Minatec Ideas Lab, CEA/Grenoble*

E-mail: christian.abry@u-grenoble3.fr

Abstract

In order to fractionate, from a behavioral and neural functional point of view, what is globally studied as the "feeling of the Beauty", we propose a brain processing distinction between: (i) core parieto-frontal multistable phenomena (e.g. the Necker cube, illustrated by Vasarely); (ii) and main prefrontal paradoxical experience (Escher cube). The latter is a stimulus that can be processed alternatively in the two different regimes.

Multistable percepts are neither limited to vision, nor to audition, but extend also to speech in the Verbal Transformation Effect (VTE: lifelifelife…–>fly). *This corresponds culturally to different traditional verbal games or pop art: just mention French* verlan *(from* l'envers*, "the reverse"). We used such speech* enactment *experiments. We were thus able to identify for the first time the neural circuit recruited by the VTE. As predicted by Reisberg, it was properly Baddeley's* Verbal Working Memory*, with two poles: Broca and the supramarginal gyrus. Then we accounted for the before unaddressed asymmetry in the VTE. The* Articulatory Loop *role was reinforced as a true* phasing control *component for speech coordinations. Hence we dubbed all this* Enactive Verbal Working Memory system *as the STABIL-LOOP, an evolutionary system for* stabilizing linguistic forms*, able to host verbal art like vision working memory systems did for pictorial art.*

1. Enactive Beauty? From philosophy to brain activity

From the Pope of deconstructionists, the late Jacques Derrida, to such a prominent enactionist as the late Francisco Varela (in the continuation of his approach, see the recent summerschool held in Fréjus, France, 6-12 sept. 2007), philosophy has remained the dominant approach to *Aesthetics*.

Symptomatic is a recent debate, launched by Krzysztof Ziarek [26], reviewing Gary Shapiro [20]. Under the heading **"**Enactive Art: Thinking and Multiplicity", he is giving the following reasons for his promotion of *enactive* as an adjective to art: "I am choosing this term "enactive" deliberately in order to emphasize two points. First, the thinking at issue here, thinking exemplified in modernist and contemporary art, is not reflective, representational, or imitative, but constitutes, instead, a kind of act. Second, this thinking is not "inactive" with the prefix "in", even though it is often mistakenly taken to be so. As homophones, **"**enactive" and "inactive" sound identical and remain indistinguishable in speech. The distinction between enactive and inactive indeed becomes visible only in writing, and its becoming visible may signal an emergence of an altogether different visibility, one on whose parameters modernity has been laboring at least since Nietzsche."

You can take this homophony as a joke or not. As phoneticians we will try to do better later. For the moment our first rebound will be to foster the well-known *enactement* procedure in *behavioral* experiments, in order to obtain *enacted* stimuli or productions: e.g. simply repeat *and* listen what you are saying. And instead of monitoring a phonetic difference between "inactive" *vs*. "enactive", we will ask you to repeat the family name of the pioneer in *Motor Phonetics*, R.H. Stetson… So you will probably stare at a *sunset*! This domain of tricks for the brain is popular in vision. Which does not mean that, even in this main science of illusions, they are fully understood: for the majority of illusions, often discovered by serendipidity, there was at the onset no theory generating facts and at the outset no theory for explaining experimentally reproducible facts. Surprisingly we will claim that for the marginal field of auditory illusions (including speech, and significantly with the most famous audio*visual* McGurk effect), and as concerns the issue of *stabilizing multistable percepts*, the task could be easier for the moment than in the very disputed corresponding arena in vision (from [11] to [3]). For this purpose we will take advantage of the famous Baddeley's working memory system, we called an “enactive memory”. Its components have been subsequently well studied in many brain imaging experiments. Among these components we fostered the *articulatory loop*, which appeared to be a *control* component for *stabilizing linguistic forms*.

Apart from mainstream philosophy of Arts, the behavioral and neural approach is now common. Let us mention Semir Zeki, founder of the *Institute of Neuroesthetics*, and other famous art lovers, neuroscientists or cognitive psychologists. They have all developed (like Varela and colleagues) more than a philosophy of *qualia*. Concerning specifically brain activity we will just mention recent experiments exploring the neural correlates of the feeling of the Beauty, a topic which became again an evolutionary issue. Lea Höffel and Thomas Jacobsen's studies addressed recently aesthetic judgement [8] and found it to be related to *moral* judgement. Their brain fMRI experiment (after [9]) evidenced mainly occipito-frontal (fronto-medial, infero-frontal) activities, but also a recruitment of the left intraparietal sulcus (LIP) for beauty as *symmetry*. We will elaborate here mainly on the *parieto-frontal* links.

2. The paradigmatic parieto-frontal enactment circuitry

At a recent summer school in Cargèse (Corsica, 4-16 june 2007) about *Conciousness & Action*, Luciano Fadiga, issued from the Parma Rizzolatti's Group, reinforced their claim that: “The classical flow-diagram describing how sensory information is processed and eventually transformed into movements by the brain has becoming more and more implausible because of neuroanatomical and neurophysiological evidence. More in detail: 1) neuroanatomical, cytoarchitectonic, histochemical and neurochemical studies indicate that the motor cortex is indeed formed by a constellation of distinct areas, *each one* ***bidirectionally connected*** *with a specific area of the parietal lobe*. 2) The neurophysiological study of *these parieto-frontal connections* suggests that they *might play a crucial role in effector-specific sensorimotor transformation*s. 3) Several motor neurons discharge also during sensory stimulation. Accordingly, ***visual stimulation modulates the activity in LIP-FEF neurons***, *objects entering the peripersonal space activate F4-VIP neurons, graspable objects and actions of other individuals visually activate canonical and mirror visuomotor neurons belonging to the F5-AIP-PF fronto-parietal circuits*.” (abstract of “Movements, actions, representations: A neural pathway to language?”).

Intraparietal areas (IP) are reciprocally connected: (i) the Lateral (LIP) with Frontal Eye Field (FEF, and SEF Supplementary EF); (ii) the Ventral (VIP) with ventral Premotor area (PMv/F4); (iii) the Anterior (AIP) and the anterior part of the inferior Parietal lobule (PF) with Broca's homologon in the monkey (F5).

From [19] and [24], the LIP-FEF connectivity has remained paradigmatic, since it emphasizes that *enacted vision* (in saccades) could be simultaneously (reciprocally in neural terms): (i) a *perceptual decision* corresponding to a stabilizing phenomenon of the object in space (LIP properties); (ii) and of a high-order (FEF and SEF are prefrontal or premotor, not simply motor eye fields) *decisional active perception*. This neurocytoarchitectonics was generalized beyond (i.e. below) LIP-FEF in [13]. Let us mention here the relevant controversy about *microsaccades*, which lasted for more than five decades, and ended recently, in the line of Hubel's team endeavor [12], by the demonstration that microsaccades prevent objects from disappearing (fading due to retinal adaptation). Note that in all levels of the visual pathway (including the cortex), neuronal firing corresponding to microsaccades increased. And there are obvious outreaches in the visual domains of binocular rivalry and illusions for perceptual stabilization.

But this does not mean that the neural experiments on multistable or impossible objects have delivered clear answers up to now. Just to mention binocular rivalry or the Necker cube, experiments and models became more and more complex (e.g. [14], [23]). And as concerns the perception and implicit memory of impossible objects, since the PET study by Daniel Schacter's group [10], which found behavioral and neural support for the *3-D Structural Description System*, for possible objects, but not for impossible ones, a recent event-related fMRI experiment concluded: “It seems more likely that a coalition of frontal and parietal brain regions involved in memory and spatial attention might be recruited for the classification of impossible objects in a manner that is highly variable from individual to individual. The discovery of such effects necessitates more specialized multivariate analysis and the utilization of subject-specific information (reaction time, recognition accuracy) in tailored design matrices” [7]. Which means that brain *perceptual* topological *decisions* (prior to reasoning) on Escher cube and the like are still poorly understood. And the same for Necker cube bistable perception, in spite of different attempts at modelling such spontaneous perception switching (for a recent *behavioral recursive nonlinear and stochastic phase oscillator model*, issued from Haken's synergetics, which integrates neurophysiological approaches and time delays, see [6]).

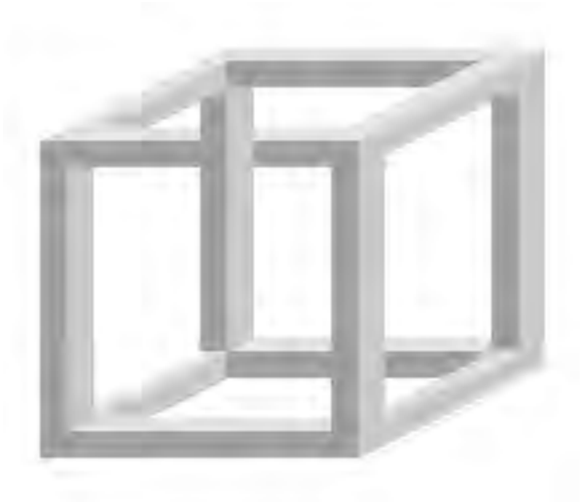


Figure 1. Escher (1958) paradoxical cube (redrawn from *Wikipedia*) can be viewed alternatively as a bistable structure. Provided you neglect (just blur your eyes) to track the (topo)logical impossibilities of this cube (*if…then* mode), you can experience it (in an *enacted* regime) as the classical Necker cube (1832): just imagine whence you enter this cube frontward? A recent study [2] revealed that tactile *exploration* of a wire Necker cube stabilizes perceptual states.

3. Enacted speech art

Well beyond this mainstream vision science of illusions, nobody would have bet to find the neural secret for what occurs when you repeat continuously "life"… and suddenly find a "fly" in your mind, which is a classical case of *enactment*. This is verbal bistability, exemplified 30 years before the Warren and Gregory's [25] *Verbal Transformation Effect*, by Stetson [22]. In fact a traditional word game for kids long before schoolyards and school institutions (see lastly French *verlan*, a special backslang: *tomber* "fall"–>*béton*). We used both nonsense syllables (see below), and as meaningful items reversible first name compounds (like *Louis-Jean–>Jean-Louis*). We decided to add two issues to the VTE [1]. First take seriously the proposal coming from Reisberg [21], that the network recruited for VTE would be the one and the same as for *verbal working memory*. Reisberg's proposal was right. In an fMRI study [15] we demonstrated for the first time the use of this circuitry for the VTE (a finding basically replicated recently by [10]). The parieto-frontal network included mainly the supramarginal (BA 40) and the inferior frontal gyri (Broca), two brain zones respectively famous for phonological word form storage and rhyme tasks (the well known rhyme effect in recall task). It must be insisted upon the evidence that just holding one item in memory for repetition does not activate this circuit. But with a perceptual bistability task in mind, this single item activates as much brainware as a "spanful" of 7-items. Extensions of BA40-Broca, testing the relationship between narrativity and numeracy are expected [5].

Our second addition was to explore the unaddressed asymmetry phenomenon: if "fly" is more reluctant to give back "life", why? Selecting properly the *articulatory loop* as the locus of a neural control bias in speech motion coordination, we were able to reject all other possible explanations of the finding that the recurrent winner-take-all of the 6 possible syllables made of one vowel (schwa) and two consonants ([p] and [s]), was [psə] [16]. Briefly said, [psə] is the most *in-phase* coordination of the vowel gesture with the consonants: the vowel is already coarticulated (coproduced) before [p] release, and [s] is ready-made to hiss within [p] (compare [əsp], the most out-of-phase as to the vowel and the consonants). This is not to say that *psi* is the optimal syllable worldwide: [p] is here obviously not acoustically salient. But from a *control* point of view, if a human being has acquired this language skill (neither English, nor Spanish), you can test it, and find that the optimum coordination is [psə], over [səp], [pəs], [spə], [əps], and [əsp], all structures attested in French phonotactics. Like you can test the optimum gait among the 3 regimes of a pony, 2 in a human, or among the many skills of a pianist, etc.

Hence we dubbed all this *Enactive Verbal Working Memory* system as the STABIL-LOOP, an evolutionary system for *stabilizing linguistic forms*. Definitively not for telephone number recall! Nor for n-load recall tasks, since as repeated by Chomsky and colleagues [4], language minimalist computation cannot count beyond 2 (there is no linguistic "rule" where a phenomenon would have to reappear each third word or third syllable, e.g. stress). That is probably why list recall studies from Baddeley's school, as noted since the 70's by his most famous student in speech perception, Chris Darwin (and as repeated recently to us by his best collaborator in the developmental field, Susan Gathercole), were finally so useless to account for *on-line* speech production/perception.

4. Finally less pictorial than speech evidence?

Why do we observe a clearer parieto-frontal network for enacted speech than for the Necker cube (not to speak of Escher)? One answer is that our STABIL-LOOP is hosted into a well known system, namely verbal working memory. Surely vision or modality-free working memory systems, including implicit memory processing [18], seem the best challenging research domains for clarifying the present issues.

Anyway we joined here a core *parieto-frontal enactivity paradigm*, with different somatotopic levels, from the eye to the mouth, i.e. different circuitry with the same enactment principle. And this could be the route for understanding perceptual decisions. What is popping out in object and art perception is limited by the same 3 seconds “refractory period”, before switching to another state of the brain. And about the same span is observed whatever the input channel, be it visual, auditory, tactile, for pictorial as well as for verbal domains.

Acknowledgments: To our colleagues working on the VTE, especially to Marc Sato.

References

[1] Abry, C., Sato, M., Schwartz, J-L., Lœvenbruck, H., and Cathiard, M-A. “Attention-based maintenance of speech forms in memory: The case of verbal transformations”, *Behavioral and Brain Sciences*, *26:6*, pp. 728-729, 2003.

[2] Bruno, N., Jacomuzzi, A., Bertamini, M., and Meyer, G. “A visual-haptic Necker cube reveals temporal constraints on intersensory merging during perceptual exploration”, *Neuropsychologia*, *45*, pp. 469-475, 2007.

[3] Carter, O., and Cavanagh, P. “Onset rivalry: Brief presentation isolates an early independent phase of perceptual competition”, *PLoS ONE*, *4*, *e343*, 5 p., 2007.

[4] Chomsky, N. *New Horizons in the Study of Language and Mind*, Cambridge University Press, Cambridge, 2000 (chap. 5: Language as a natural object).

[5] Diaferia, M.-L., and Abry, C. “Du conte au compte: pour une ethnographie du développement de la narration et de la numération”, *DECOLAGE III, Le Développement Conceptuel et Langagier de l’Enfant*, (Reims 23-24 Juin 2005) http://www.univ-reims.fr/Labos/Accolade, 2005.

[6] Fürstenau, N., “Modelling and simulation of spontaneous perception switching with ambiguous visual stimuli in augmented vision systems”, *Perception and Interactive Technologies*, LNAI 4021, pp. 20-31, Springer, Heidelberg, Berlin, 2006.

[7] Habeck, C., Hilton, H.J., Zarahn, E., Brown, T., and Stern, Y. “An event-related fMRI study of the neural networks underlying repetition suppression and reaction time priming in implicit visual memory”, *Brain Research*, *1075, 1:23*, , pp. 133-141, 2006.

[8] Jacobsen, T., Schubotz , R.I., Höfel, L., and Cramon, D.Y.v. “Brain correlates of aesthetic judgment of beauty”, *NeuroImage*, *29*, pp. 276-285, 2006.

[9] Kawabata, H., and Zeki, S. “Neural correlates of Beauty”, *Journal of Neurophysiology*, *91*, pp. 1699–1705, 2004.

[10] Kondo, H.M., and Kashino, M. “Neural mechanisms of auditory awareness underlying verbal transformations”, *NeuroImage*, *36*, pp. 123-130, 2007.

[11] Leopold, D.A., Wilke, M., Maier, A., and Logothetis, N.K. “Stable perception of visually ambiguous patterns”, *Nature Neurosciences*, *5:6*, pp. 605-609, 2002.

[12] Martinez-Conde, S., Macknik, S.L., Troncoso, X.G., and Dyar, T.A. “Microsaccades counteract visual fading during fixation”, *Neuron*, *49*, 2006, pp. 297-305, 2006.

[13] Matelli, M., and Luppino, G. “Parietofrontal circuits for action and space perception in the macaque monkey” *NeuroImage*, *14*, pp. 27-32, 2001.

[14] Meng, M., and Tong, F. “Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures”, *Journal of Vis*ion, *4*, pp. 539-551, 2004.

[15] Sato, M., Baciu, M., Lœvenbruck, H., Schwartz, J-L., Cathiard, M-A., Segebarth, C., and Abry, C. “Multistable representation of speech forms: A functional MRI study of verbal transformations”, *NeuroImage*, *23*, pp. 1143-1151, 2004.

[16] Sato, M., Schwartz, J-L., Cathiard, M-A., Abry, C., and Lœvenbruck, H. “Multistable syllables as enacted percept: A source of an asymmetric bias in the verbal transformation effect”, *Perception & Psychophysics*, *68:3*, pp. 458-474, 2006.

[17] Schacter, D.L., Reiman, E., Uecker, A., Polster, M.R., Yun, L.S., and Cooper, L.A. “Brain regions associated with retrieval of structurally coherent visual information”, *Nature*, *17:376(6541)*, pp. 587-590, 1995.

[18] Schacter, D.L., Wig, G.S., and Stevens, W.D. “Reductions in cortical activity during priming”, *Current Opinion in Neurobiology*”, *17*, pp. 171–176, 2007.

[19] Schall, J.D., Morel, A., King, D.J., and Bullier, J. “Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams”, *Journal of Neuroscience*, *15*, pp. 4464–4487, 1995.

[20] Shapiro, G. *Archaeologies of Vision: Foucault and Nietzsche on Seeing and Saying*, University of Chicago Press, Chicago, 2003.

[21] Smith, J.D., Reisberg, D., and Wilson, M. “The role of subvocalization in auditory imagery”, *Neuropsychologia*, *11*, pp. 1433-1454, 1995.

[22] Stetson, R.H. *Motor phonetics. A study of speech movements in action*, in *Archives néerlandaises de phonétique expérimentale*, *t.3*, 1928.

[23] Tong, F., Meng, M., and Blake, R. “Neural bases of binocular rivalry”, *Trends in Cognitive Sciences*, *10:11*, pp. 502-511, 2006.

[24] Umeno, M.M., and Goldberg, M.E. “Spatial processing in the monkey frontal eye field. I. Predictive visual responses”, *Journal of Neurophysiology*, *78*, pp. 1373–1383, 1997.

[25] Warren, M.R., and Gregory, R.L. “An auditory analogue of the visual reversible figure”, *American Journal of Psychology*, *71*, pp. 612-613, 1958.

[26] Ziarek, K. “Enactive Art: Thinking and Multiplicity”, review of Shapiro, G. [20], *Bryn Mawr Review of Comparative Literature*, *6:1*, Winter 2007.

On-line physically based control of mass-interaction models

Ali Allaoui* Matthieu Evrard* Annie Luciani*
(*)ICA Laboratory, France
ali.allaoui@imag.fr, matthieu.evrard@imag.fr, annie.luciani@imag.fr

Abstract

Physical modeling allows producing a large set of motions for computer-based animation. The mass-interaction formalism is an interesting way to design physically based models (modularity, simplicity of each module, genericity) and it leads to produce different behaviors: smoke, paste, liquid, crowd... A specific difficulty is to deal with phenomena that are highly non-linear by nature. The point of this article is to present a way of dealing with such phenomena by on-line dynamic changes of the physical parameters, leading to non-linear dynamic behaviors. We describe a novel method inside the mass-interaction formalism that permits to physically control such dynamic parameters changes during the simulation of the physical model. The method consists in introducing in the modeling process, a "controller component" layer, also based on the mass-interaction formalism, that is in charge of the on-line modifications of the concerned physical parameters through a "parameter control module" (PCM). The article then illustrates the method on one example: a usual human gesture, a repeated jump, which requires modifying in a cyclic manner the tonicity of the muscle, during the jumps.

1. Introduction

Physical modeling has become a widely used means to design animation. The mass-interaction system is able to produce a wide variety of different behaviors [5][7][9][12]. Very soon, it has been improved to obtain complex evolving phenomena by introducing an on-line control of some physical parameters, such as the rest length [2][6][8] or the stiffness of springs [10]. The work presented here proposes a control method entirely based on mass-interaction formalism, able to modify on-line the physical parameters of the modeled object by a physically based process, in order to create physical coherent non-linear behaviors. This method is based on three components:

(1) A generic module, called "parameter control module" (PCM), for the on-line control of physical parameters of a 3D model, compatible with the mass-interaction formalism

(2) A 1D mass-interaction network representing the controller disposal that provides scalar to (1)

(3) Interaction between the 3D model and (2) that provides information on the 3D model state to (2).

In section 2, this approach is compared to other implementations of controls of physically based modules. Section 3 deals with the implementation of our method within our user-friendly modeling MIMESIS software for mass-interaction model design [9]. The method is illustrated by an exemplary modeling case, a usual human gesture, a repeated jump, which requires modifying in a cyclic manner the tonicity of the muscle.

2. Global context

Physical modeling is a generative way of producing motion. As in Physics, physical *parameters* are data such as stiffness, viscosity, thresholds in distance and velocity, static and dynamic coefficient of friction, inertia, while positions, velocities, angles, forces, torques, … are inputs or outputs *variables* of the physical model.

When dealing with the control of physically based models, we can distinguish two main directions: (1) one aiming at acting on the variables, for example forces, positions, velocity fields as done in [1] and [14]; (2) a second acting on the physical parameters of the 3D objects of the scene. In this paper, we focus on this second type of control.

The main way in computer graphics for dealing with this type of control is to specify high-level objectives in order to synthesize controllers by means of optimization methods.

Our contribution aims at considering the physical model and the controller as a whole dynamical system, designed by means of the same unified formalism.

Hence, we propose to extend the possibilities of the mass-interaction formalism as described in figure 1 with our controller system in two layers:

(1) The controller itself: it is a one-dimensional mass-interaction system. The controller is in physical interaction with the 3D object. It reacts on line to the dynamical states of the 3D objects in their physical environments.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

(2) A set of modules, one by parameter controlled, that modify the physical parameters of the 3D objects from the data provided by the controller.

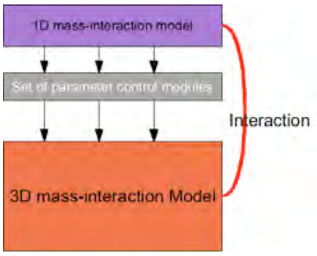


Figure 1: On-line physically based control of physical parameters

So doing, we make possible what can be called non-linear parameters auto-regulations as it is in several complex non-linear phenomena like Van der Pol oscillators for example [13].

Differently than all the previous approaches, our method allows the user to build the controller. By using the physical mass-interaction paradigm, the user designs a whole closed loop dynamical system with a simple and modular language. Actually, one can figure out easily what the system represents, because its parameters (elasticity, viscosity) are meaningful in everyday life. Of course, this simplicity is relative, as it goes more and more complex when the number of modules increases.

3. The parameter control module (PCM) in MIMESIS software

The parameter control module takes place within a user-friendly modeling MIMESIS software [3]. It is based on mass-interaction modeling ([5]). In MIMESIS, the mass-interaction formalism is the core of the creation process at hand. It is therefore able to manage a lot of modules, edit initial conditions and physical parameters, simulate the designed model, import and export the signals representing variables that evolve in time, coat the obtained movements, etc.

Until now, MIMESIS software has support for two categories of modules, the MATs and the LIAs (figure 2, up). A MAT module is a module that takes forces as in input and returns a position. It is a kind of material element which main parameter is its mass (or inertia). A LIA module connects two MATs. It represents the interaction between both MATs. This interaction can be a simple linear damped spring but also a more complex non-linear interaction such as cohesion, plasticity, friction, etc. Physical parameters of LIAs are stiffness, viscosity, thresholds on distance, static and dynamic coefficient of friction, etc..

This work introduces a new module called PCM (parameter control module). It allows modifying a physical parameter of a MAT or a LIA by the position of a MAT, i.e. the value of the parameter will depend on the position of this MAT. Unlike MAT and LIA, that are bidirectional modules (figure 5), it is an oriented module: it takes a position from the MAT and computes the parameter value through a mathematical function. As a parameter is a scalar value, the simplest choice for the MAT that modifies this parameter through the PCM module is a 1D MAT, i.e. a MAT that produces a 1D scalar data. (see figure 2, down).

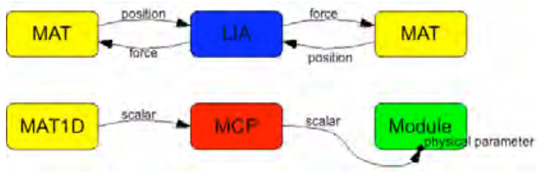


Figure 2: Modules in MIMESIS software

4. Modeling an elementary human motor gesture: jumps

4.1. The phenomenon and the model

A repeated jump requires to modify in a cyclic manner the tonicity of the muscle, during the jumps. It is a basic case of non-linear dynamic behavior. In this paper, we are interested in introducing this type of modification. We used a simple passive mass-interaction model of a jump previously introduced by Hsieh in [4]. In this model, the muscles of the leg are approximated as a unique spring connecting the foot and the pelvis (figure 3) and a jump is triggered by the initial velocity of the pelvis. In order to render a jumper who will jump repeatedly, i.e. to start a new jump as soon as he has landed, it is necessary to make this model active, for example by changing its inner dynamic state during the simulation. It may be done by changing the physical parameters according to the state variables.

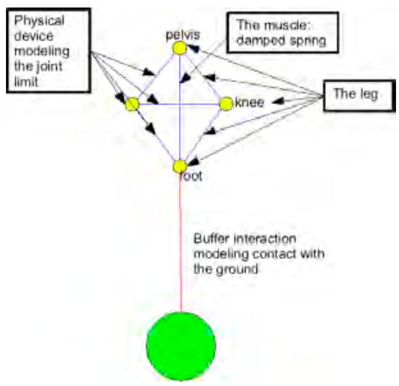


Figure 3: 3D model of a jumping leg

4.2. The controller

A way to do this is to modulate the rest length of the muscle by means of a mass-interaction controller described in figure 4.

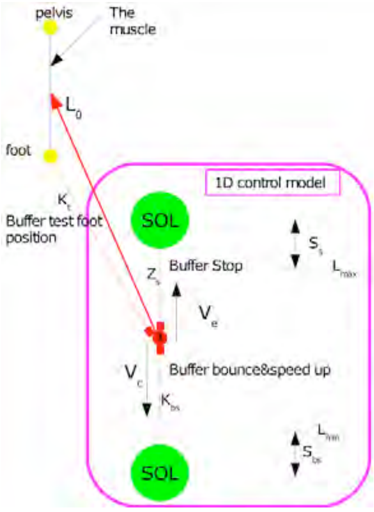


Figure 4: Repeated jump model and its control

The position of the mass n°1 (see figure 4) controls the muscle rest length through a PCM module (the arrow pointing to the muscle).

The jumper performs his jumps in a repetition of 4-steps cycle :

1. The jumper prepares a jump by contracting his muscle. In the controller view, mass n°1 moves from L_{max} to L_{min}.

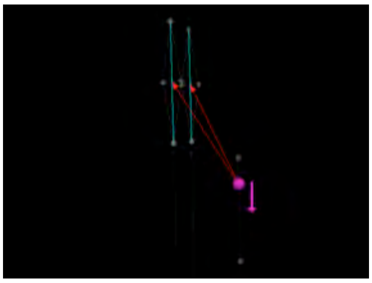


Figure 5: Step 1

2. He extends his muscle quickly. Mass n°1 bounces on the fixed mass located at L_{min}, and speeds up thanks to a well-tuned elastic buffer.

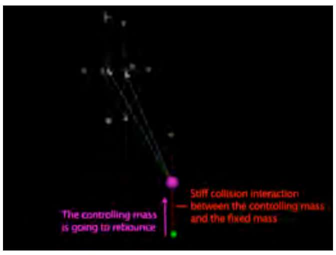


Figure 6: Step 2

3. During the take-off, when the mass n°1 reaches L_{max}, it is stopped by a highly viscous buffer, leading to maintain the legs in extension during the jump.

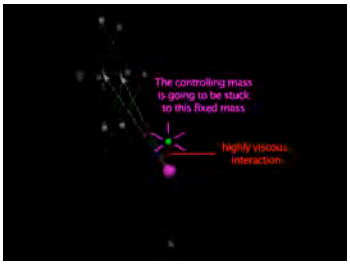


Figure 7: Step 3

4. During the landing, when the jumper hits the ground, he also hits mass n°1, so the jumper is triggering a new jump (step 1).

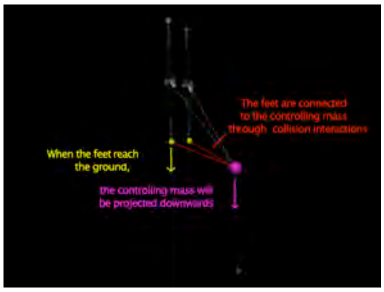


Figure 8: Step 4

4.3. Tuning the controller model

In our model, the state variables of the current jump (cycle n-1) that are relevant in the performance of next jump (cycle n) are: the height of the jump H_j, the velocity of contraction V_c, and the velocity of extension V_e.

Moreover, these 3 parameters are influencing each other, from a cycle n-1 to the next cycle n:

$H_j(n)=f(V_e(n),V_c(n))$

$V_c(n)=g(H_j(n-1))$

$V_e(n)=h(V_c(n))$

The functions f, g, h can be tuned implicitly, i.e. without having an explicit expression of them, by tuning the three parameters of the buffer interactions (buffer stop, buffer bounce and speed up, buffer test foot position, in figure 4). First, the stop buffer is just a viscous interaction, tuned to stop the agitator the closest to L_{max}. Then, the buffer "test foot position" is the softest possible to detach the agitator from the influence of the stop buffer, so that V_c is the smallest possible. Finally, the buffer "bounce & speed up" takes advantage of a discretization artifact, that can be usefully mastered thanks to [11] which makes the correlation between stiffness values and the gain or loss of energy according to the used discretization scheme.

5. Conclusion

Generating highly non-linear movements is always a tricky task in modeling. However, to obtain a certain degree of complexity, non-linearities must be taken into account. Indeed, a lot of natural well-known phenomena are themselves deeply non-linear, as those in which there are changes in the physical states of the matter (for example, fluid - solid transitions), modulation of physical parameters according to the states of the system (for example, modulation of the elasticity according to the elongation of the spring), or in active systems such as human locomotion. After the huge developments in computer animation, we assumed that it is a way to reach a new step in the complexity and the expressiveness of synthetic motions.

Two questions have been risen in the work presented here: first, how to give access to the user in the designing of such type of control? And second how implement it? We proposed to introduce a specific module for controlling physical parameters in a mass-interaction generic system and we set up methods to design controllers that are also physically based and consequently in closed dynamic interaction with the objects to be controlled. So doing, we initiated the process of implementation of inmost closed-loop control systems, as it is necessary in complex non-linear systems. This method corresponds really to a shift in the way of thinking about the control process: from explicit control to implicit dynamic control. We experimented this approach in two complementary examples: the modulation of elasticity and the modification of the tonicity of a simple muscle-like object as it occurs in an active system.

The modules and models developed in this work have been implemented for the first time, within a user-friendly interface, which allow end-users to design mass-interaction models. Consequently, they are available for any user who wants to understand the method and so doing, to become able to use it to design his own physically based models including dynamic parameters control. We assume that it could be the best "practical" way to introduce them freely in the domain of complex dynamic modeling and to constitute their own expertise by themselves.

References

[1] R. Barzel, A. H. Barr. A modeling system based on dynamic constraints. In Proceedings of SIGGRAPH 1988, ACM SIGGRAPH, 1988, pp. 179-188.

[2] P. Boulenguez, A. Luciani. A Physical Particle Model of Morphogenesis. WSCG'03 Plzen (Czech Republic).

[3] M. Evrard, A. Luciani, N. Castagné. MIMESIS : Interactive Interface for Mass-Interaction Modeling, In Proceedings of CASA 2006, Geneva, July 2006, Nadia Magnenat-Thalmann & al. editors. pp177-186.

[4] C. Hsieh, A. Luciani. Generating Dance Verbs and Assisting Computer Choreography. *Proceedings of MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pp 774—782. 2005.

[5] A. Luciani, S. Jimenez, J.-L. Florens, C. Cadoz, O. Raoult. Computational physics: a modeler simulator for animated physical objects. *Proceedings of the European Computer Graphics Conference and Exhibition*. Eurographics'91, pp 425-436, Vienna, Austria, Elsevier Ed, Sep. 1991.

[6] G.S.P Miller. The motion dynamics of snakes and worms. *Computer Graphics*, vol. 22, SIGGRAPH'88, 169-178, Aug 1988.

[7] G.S.P. Miller, A. Pearce. Globular dynamics : a connected particle system for animating viscous fluids. *Computers & Graphics*, vol. 23(3):169-178, Elsevier, 1989.

[8] Sodaplay java applet - http://www.sodaplay.com

[9] X. Provot. Deformation constraints in a mass-spring model to describe rigid cloth behaviour. *Proceedings of Graphics Interface 1995*,Graphic Interface'95, pp 147-154, Quebec City, Canada, (May 1995).

[10] D. Terzopoulos, J. Platt, K. Fleisher. Heating and melting deformable models : from goop to glop. *Graphics Interface'89*, 219-226, 1989.

[11] F. Thil. Simulations de collisions et application à la synthèse d'objets sonores et percussifs. Rapport de stage de DEA, ACROE-ICA Grenoble 2003.

[12] D. Tonnesen. Modeling liquids and solids using thermal particles. *Graphics Interface'91*, 255-262, 1991.

[13] B. van der Pol, J. van der Mark. Frequency demultiplication. *Nature*, 120, 363-364, (1927)

[14] A. Witkin, M. Kass. Spacetime constraints. *Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, p.159-168, June 1988

Modelling Enactive Interaction with a perceptual supplementation device

Yusr AMAMOU John STEWART
COSTECH, Technological University of Compiègne, France
yusr.amamou@utc.fr & john.stewart@utc.fr

Abstract

"Enactive knowledge" is distributed across all the interactions between an organism and its environment. When a human subject interacts with a computerized virtual environment, his motor acts determine sensory feedback from the machine, giving rise to sensory-motor dynamics.

The traces of these interactions, which are readily retrieved from the computer, complete information concerning the user's activities. The analysis of traces makes it possible to describe the sensory-motor dynamics, and to characterize the variety of strategies employed by the users.

1. Introduction

The devices of sensory substitution, also called supplementation perceptive devices [7], make it possible to give an access to objects generally reserved for one sensory modality, such as vision, via another sensory modality, such as touch. These prosthetic devices were developed for assistance to blind people since the end of the Sixties [3, 5]. A camera controlled by the subject and a tactile matrix put on the skin which translates the image collected by the camera into a form "in relief". After a short training, and in spite of the poverty of tactile "information", the subjects manage to recognize and locate objects and people [4].

If the subject does not have means of moving the camera, which for example is fixed on the table, it does not have percept. It is by establishing correspondence between actions and sensory returns that the subject builds invariants. The subject forgets the tool itself which is in contact with its skin to perceive the remote object [4, 6]. The explanation is thus in the sensory-motor approach of perception.

O'Regan and Noë [9] define the distinction between sensory modalities according to the criterion of sensory-motor contingencies. The different phenomenal "feels" and the specific characteristics of each perceptual modality are explained in terms of dynamics structures. What will differentiate vision from another sensory modality is the structure of the laws controlling the sensory changes produced by the various motor actions, i.e. the sensory-motor contingencies constraining the visual exploration.

Because the sensory-motor contingencies of each modality have different invariant properties, the structure of the laws which control perception is different according to each modality.

Our work is situated within the framework of a sensory-motor theory of perception [9, 10, and 11] according which information on the world is contained in the properties of the structural invariants binding the actions to resulting sensory stimulations. What is perceived and recognized are not invariants of the stimulation, but invariants of sensory-motor loops inseparable from the subject's activity (see Figure 1). It is by means of their actions that the subjects isolate the regularities in the relation between their motor actions and the sensory changes which result from it. It is by means of their actions that the subjects build the covariation laws between the motor outputs and the sensory inputs that result.

According to these theories, a sensory experience, such as visual experience, is a mode of activity implying a practical knowledge of the possible behaviors and associated sensory effects.

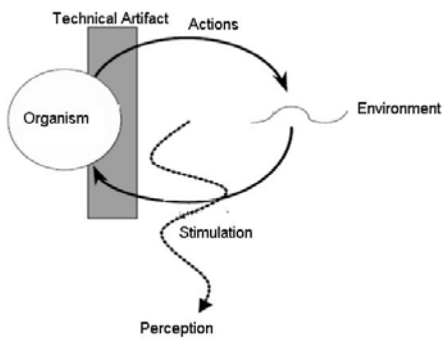


Figure 1. Sensory-motor coupling between organism and its environment

According to this perspective, we propose that the first step of tool appropriation, the contact stage, consists of extracting the regularities between the actions carried out in organic space and resulting sensory stimulations (see Figure 1). The users thus learn the sensory-motor regularities necessary to stabilize and preserve the perceptual contact with the stimulus.

Thus, by dint of their actions, new users of a coupling device with the environment have the possibility to extract the invariant relations between the movements carried out and the resulting sensory stimulations. It is on this basis that a perception via the device can be built.

The perceptual supplementation technologies allow a study of the genesis of perception. They enable us to demonstrate the role of movement and user engagement in the process of perceptual constitution via the interface. The sensory returns, by the device, do not give access to the totality of the graphic form, but they oblige the coupling and support the invariance.

The user sees the environment differently than the designer. He does not use the same vocabulary as the designer and explores his environment differently. The traces produced by user are digital indications of knowledge correlated with the user's actions.

The analysis of the traces makes it possible to seek the similar cases by analogical research and makes it possible to detect the sometimes unexpected errors made in the interaction. One of the objectives of the analyses of the traces is the training of users of complex technical artefacts. After descriptive analysis of trajectories we elaborate a model of the strategies which generate the trajectories. In this study, we examine the precision and the robustness of the strategies used by subjects to identify shapes (simple graphic forms).

2. Experimental device

The actions of the subject (blind or blindfolded) consist in moving the stylus on a graphic tablet (see Figure 2). These movements command the position of the cursor on a computer screen. When the cursor receptor field controlled by the stylus crosses the virtual form, the computer treats the signals and transmits to the stimulators, the Braille cells. The subject has the finger placed on the tactile stimulators, thus he receives a tactile feedback. The only way the subject can perceive the form is by moving the stylus so as to explore the contours of the virtual form (the form exists on the screen and the subject explores it via movements on the tablet).

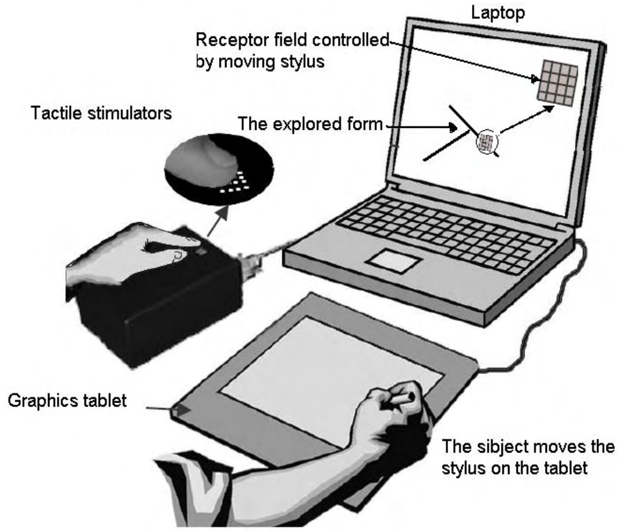


Figure 2. Tactos device

An example of a "trajectory of perceptual exploration" is shown in Figure 3. This subject explored and identified a shape composed of two simple segments forming the letter T.

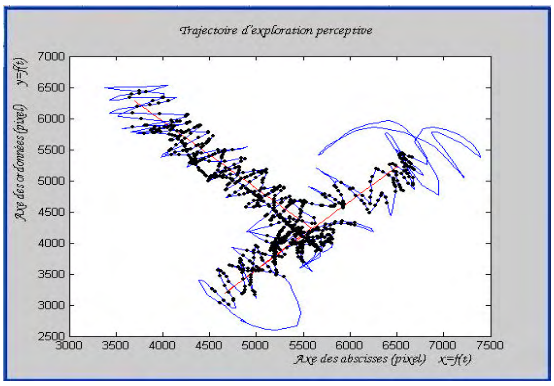


Figure 3. Trajectory of perceptive exploration, strategy used "Micro sweep"

The minimalist supplementation device forces a maximum externalization of the perceptive activity [6, 8], i.e. spatial and temporal deployment, which allows a systematic study of the dynamics perceptive trajectories and strategies produced.

The subject (blind or blindfolded) used "Tactos" device for first time, in this experiment the subject uses the sensory feedback to guide subsequent actions and the last determine subsequent the sensory feed back; together, this constitutes a dynamics of sensory-motor loop. An analysis of exploratory trajectories suggests that the subjects adopted minimum one of the different strategies of form perception. 34% of subject used the strategy of "Micro-sweep" (strategy used on figure 3). The subject deliberately implements an oscillatory movement while sweeping around the form [1].

We defined descriptors which categorize these strategies such as the Fourier transform. We could show that the values of descriptors applied to the perceptive trajectories generated automatically by a simulator answer the criteria established to describe and differentiate the strategies of the human subjects [2].

3. Procedure and results

The perceptive trajectories are reproduced from the recorded data files and are filtered in order to be homogeneous. A preceding study revealed that the same user, even within same exploration, can change strategy [2].

The traces that we used for our analysis and modelling trajectories contain information about the efficient subject's strategy.

Our final objective is to find a model which identifies the strategy employed for each given trajectory.

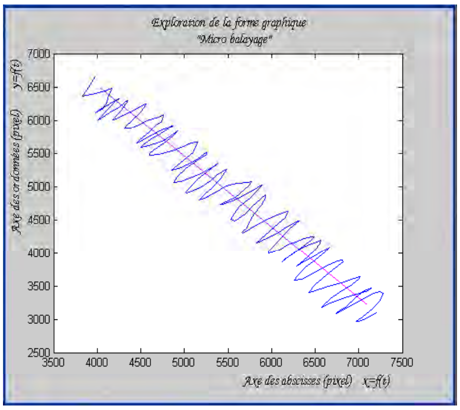


Figure 4. A micro sweep trajectory

The first step was an examination of the different trajectories, and a qualitative identification of the various strategies used by the subjects.

The second step was to establish a list of quantitative descriptors potentially able to identify these different strategies. Three descriptors appeared to be particularly relevant:

D_1: the proportion of power in the high-frequency component (>0.5 cycle per second) of the Fourier analysis.

D_2: the proportion filtered of power in the high-frequency component (>0.5 cycle per second) of the Fourier analysis and

D_3: the mean value of the oscillation amplitude.

The third step was to attribute a score, between 0 and 100, expressing the degree to which a given trajectory belonged to a particular type of strategy. In this paper we concentrate on the "micro-sweep" strategy (Figure 4). These scores were attributed, subjectively, by two independent experts[1] (Table 1).

The fourth step was to calculate a linear regression to predict the scores on the basis of the input values (D1 and D2 and D3):

$a_0+a_1D_1+a_2D_2+a_3D_3$, with: a_0 =-2.78, a_1=47.71, a_2=21.82 and a_3=22.22.

Table 1. A portion of data Table.

Trajectory	D1	D2	D3	Score
T1	0,99	1,00	1,24	100
T2	0,99	1,00	1,13	100
T3	0,63	0,99	1,01	0
T4	0,60	1,00	1,07	0
T6	0,22	0,80	0,40	0
T7	0,52	0,85	0,54	25
T8	0,60	0,80	0,93	18
T9	0,15	0,61	1,24	30

To validate this model we must study its precision and robustness. The precision is defined by the coefficient of correlation between the observed scores (Table 1) and those predicted by the model. The value obtained, 0.83, is close to 1, i.e. the predicted values are near the values observed by the experts.

However, the precision of the model is not enough; to evaluate its robustness we used a Monte-Carlo method. Thus, we extracted a random 10% of input data; applied the same method of linear regression to the remaining 90% rest of data; and tested the accuracy of predictions for the 10% extracted.

We repeat this process 500 times (Figure 5). We save the predicted values and the observed values on different matrix each time we generated a model. We will have 500 different models generated with same method of linear regression, the difference are in the coefficient a_0, a_1, a_2 and a_3.

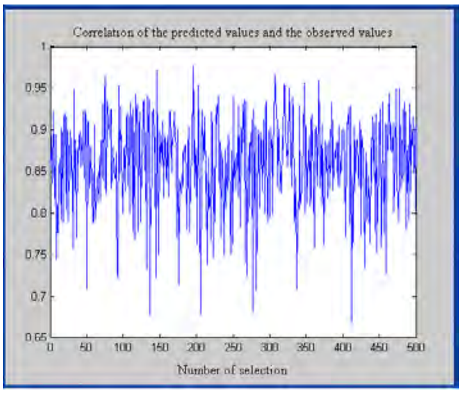


Figure 5. Coefficient of correlation between the predicted values and the observed values

[1] The degree of inter-expert agreement was high, correlation >0,9.

After that we calculate the coefficient of correlation (to study the robustness of the method) between the predicted values and the observed values. The result was 0.82.

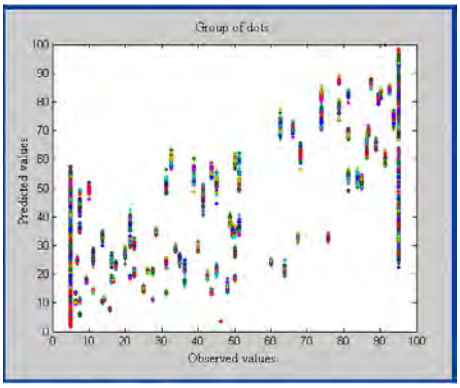


Figure 6. Group of dots of the observed values in function of the predicts values

We present in figure 6 the output real (observed) values in function of the predicted values (with the model generated).

4. Discussion and conclusions

The subjects are novices and they adopted alone these perceptual strategies. These strategies are proven effective for the form identification. According to this study the appropriation of device, we elaborated a model to be implemented and to be an automatic support to evaluate which strategy was used by subject.

The subject have a difficulty to explain what did he do, the strategy modelling made the subject seem what did he do, and he can improve his using for better adoption and appropriation of device via this efficient strategies.

So the models of strategy can be an important help for subject and he can study his performance and his evolution strategies. This model will help subject to reduce the time of the phase of familiarization of device before the phase of identification.

For that, we study the robustness and precision of the model, we must be sure that our model is pertinent. Be sure that model was not learning from our data and when this model will be applied to new data have not a big difference between the predicted values and the real values. The result was positive and we can implement this model in our device, and help the new user to appropriate the new device with the automatic feed back model.

References

[1] Amamou, Y. and Stewart J. Perceptive strategies with an enactive interface. *The Third International Conference On Enactive Interfaces*, (2006), 101-102.

[2] Amamou, Y. and Stewart J. Analyse comparative de trajectoires d'exploration perceptive. *Third Meetings Young Researchers in IHM ,* 2006.

[3] Bach-y-Rita, P., Collins, C. C., Saunders, F., White, B., and Scadden, L. Visual substitution by tactile image projection. Nature, 221, (1969). 963-964.

[4] Bach-y-Rita, P. (1972). Brain Mechanisms in Sensory Substitution. New York: Academic. Press.

[5] Kay, L. An ultrasonic sensing probe as a mobility aid for the Blind. Ultrasonics, (1964), 2, 53.

[6] Lenay C., Canu, S. and Villon, P. Technology and perception: the contribution of sensory substitution systems. In Proceedings of the Second International Conference on Cognitive Technology (1997), (pp. 44-53).

[7] Lenay C., Gapenne O., Hanneton S., Marque C. et Genouëlle C. (2000) La suppléance sensorielle : Limites et perspectives. In Y. Hatwell, A. Streri & E. Gentaz (Eds), Toucher pour Connaître. Psychologie Cognitive de la Perception Tactile Manuelle, Paris, PUF, pp. 287-306.

[8] Lenay, C., Gapenne, O., Hanneton, S., Marque, C., & Genouëlle, C. Sensory substitution: Limits and perspectives. In Y. Hatwell, A. Streri & E. Gentaz (Eds.), Touching for Knowing, (2003), (275-292): John Benjamins.

[9] O'Regan K.J. and Noë A. A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences 24, (2001), 5-115.

[10] Gibson, J. The Senses Considered as Perceptual Systems, (1966).

[11] Piaget, J. La construction du réel chez l'enfant. Neuchatel: Delachaux & Niestlé, (1937).

Towards Tangible Enactive-Interfaces

Peter Bennett Sile O'Modhrain
Sonic Arts Research Centre, Queens University, Belfast
E-mail: pbennett02@qub.ac.uk, sile@qub.ac.uk

Abstract

This paper compares the disciplines of Enactive Interfaces and Tangible User Interface design with the aim of showing that by considering the design knowledge of Tangible User Interfaces with the cognitive depth of the theories of enaction, a new type of "enactive tangible-user-interface" can be designed. Introductions to both disciplines are given, before a method of gauging the enactive potential of particular TUI's is described. Finally a discussion of the two disciplines' combination is given with such questions as, "are all tangible-user-interface's enactive?" and "are all enactive interfaces tangible?"

1 Introduction

Tangible User Interface (TUI) design is now a well-established field in comparison with the study of Enactive Interface's (EI's) which is relatively young. It is proposed here that it is possible to use the knowledge of TUI design for the benefit of developing EI's. Introductions to both TUI's and EI's are given in the following sections.

2 Tangible User Interfaces

TUI's, in Ullmer and Ishii's definition, are interfaces that "give physical form to digital information, employing physical artefacts both as representations and controls for computational media"[16]. Several frameworks for TUI design have been proposed, including the MCRpd model by Ullmer & Ishii [16] and the embodiment and metaphor based taxonomy of Fishkin [7] which will be discussed in this paper. First, as an introduction to the different types of TUI, four classifications of tangible interface instance are outlined below.

2.1 Tangible Interface Instances

In [16] Ishii identifies four general types of TUI with the categories: Spatial, Constructive, Relational and Associative. Spatial systems rely on the spatial interpretation of multiple physical objects within a common frame of reference. A typical case is the use of tangible blocks on a horizontal augmented surface, an example being the AudioPad system [13]. Constructive systems rely on the building-block principle, and involve using a number of objects that can be combined together, such as in the Blockjam system [12]. Relational systems couple digital information with a physical object, for example of the slot-machine programming system [14] that allows users to program a Logo 'Turtle' through the use of cards that represent commands. Associative systems are similar to relational systems with the difference that each physical-digital link is fixed, and there is no interaction between the objects. A good example of this type of system is the music-bottle project [9] in which individual bottles are linked with a sound that can be played by removing the bottles lid. A sub-category is defined for devices that are both Constructive and Relational. An example from this system is the bricks tangible programming project [10] as it relies both on linking digital information in a dynamic manner and on the constructive arrangement of its blocks.

2.2 Embodiment and Metaphor Based Taxonomy

A taxonomy has been developed by Fishkin [7] that uses the concepts of embodiment and metaphor to classify TUI's. The rationale of Fishkins's taxonomy is that tangible interfaces are a particularly broad area of research, and a simple binary definition to decide whether an interface is tangible is not sufficient. The solution is to create a two dimensional taxonomy that allows for a scale of tangibility. The two axes of this taxonomy are embodiment and metaphor.

Fishkin's definition of embodiment is effectively a measure of how close the digital output is to the input, and also to what extent the user thinks that the states of the system are inside the device. The scale is defined as ranging from 'Distant' (the output is removed from the input) through 'Environmental' (output is around the user) to 'Nearby' (the output is near to the input), ending up with 'Full' (the output device is the input device).

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

The scale of Metaphor looks at how the use of the interface can relate through metaphor to a real-world concept. The scale ranges from 'No Metaphor' (no analogy to an existing object or action), to 'Noun or Verb' (the interface either looks like (noun) or acts like (verb) something else), to 'Verb and Noun' (the interface both looks and acts in a similar way to another object), finally to 'Full' (the user does not need to make an analogy because there seems to be no distinction between the virtual and physical systems).

Looking at the utility of this taxonomy, it is shown how it can incorporate Holmquist's theory of containers (objects to move digital information), tokens (icon-like objects used to access stored information) and tools (used to manipulate digital information)[8] as well as the 'object as' theory [17]. Of particular interest to this paper is the study that looked at how tangible interfaces have evolved within a particular task domain. It was found that in the three areas in which there had been multiple projects (children's storytelling, tangible workbenches and control widgets on an augmented desktop) the evolution of the field had progressed from the 'no-metaphor/distant-embodiment' corner of the taxonomy to the 'full-metaphor/full-embodiment' corner. It is proposed that this corner of the taxonomy is the one most likely to hold enactive-TUI's, as full-metaphor and full-embodiment are conducive to the transfer of enactive knowledge.

3 Enactive Interfaces

Enactive interfaces are a classification of interface that allow the expression and transmission of enactive knowledge. Enactive knowledge, as opposed to symbolic or iconic knowledge, is a form of knowledge that is stored in bodily sensori-motor responses. The handling of enactive knowledge via the means of an enactive interface can be considered a particularly direct means of communication between human and computer. Enactive interfaces are desirable because they allow the user to utilise their pre-conceived knowledge of interacting with the world when using the interface.

3.1 Criteria of Embodied Interaction

In his development of a theory for Enactive Instruments [1] Armstrong outlines a list of criteria for embodied interaction. These are summarised below:

1. Embodied activity is situated. The agent is situated in an environment.
2. Embodied activity is timely. Real-word activity requires real-time constraints.
3. Embodied activity is multimodal. Concurrent use of multiple sensory modalities with the possibility of cross coupling between the modalities.
4. Embodied activity is engaging. The agent is required by the system and is actively engaged with it.
5. The sense of embodiment is an emergent phenomenon.

Here we will only be considering the first four criteria as possible design aids, as emergent phenomena are something that can be seen to be a result of an embodied system, rather than something that can be designed for.

4 Comparison of TUI's and EI's

TUI's and EI's can be compared on two different levels, the first is to consider their general similarities and dissimilarities, the second is to make a deeper theoretical comparison using the criteria of embodied interaction.

4.1 Similarities

Both TUI's and EI's have a reliance on the haptic modalities. TUI's inherently have a haptic element due to their tangibility; even if the haptic modalities are not used for active sensory display, then they will at least be present in a passive manner for input. EI's are reliant on haptics in a slightly more subtle way as they deal in the transmission and storage of enactive knowledge, which is based on motor skills which are intimately linked with haptics.

TUI's and EI's are both multimodal in general. Multimodality is the third criteria in Armstrong's definition for embodied interaction so it can be assumed that EI's that support embodied interaction will also need to be multimodal. TUI's do not necessarily need to be multimodal, however in practice a large number of them are. A reason for this can be seen in the way that the standard haptic interaction of a TUI is usually augmented with visual and audio feedback thus creating a multimodal system.

4.2 Dissimilarities

The main difference between TUI's and EI's is in the form that they represent knowledge. Bruner [4] describes the three possible types of knowledge used when interacting with the world as symbolic, iconic and enactive. Symbolic knowledge involves conceptualization and abstract reasoning, iconic knowledge involves visual recognition and the ability to compare and contrast, and enactive knowledge is constructed on motor skills. Enactive interfaces work on the enactive level of knowledge, whereas TUI's are free to work within any of the modes. Because they are most often comprised of objects contextualised in a physical environment, TUI's inevitably rely to a great extent on the enactive level, however they tend to normally operate in the iconic or symbolic realms. An example of an iconic TUI is the 'Urp' urban planning system in which the tangible element takes the form of the buildings they represent. An example of a symbolic TUI is the Audiopad [13] which is designed for realtime music performance. The user manipulates radio-frequency tagged pucks around a tabletop, navigating interactive text-based branching menus that are projected onto the table's surface.

4.3 Comparison using Embodied Interaction Criteria

Using Armstrong's criteria of embodied interaction, as outlined in section 3.1, it is possible to see which of the criteria is applicable to TUI's.

Starting with the criterion 'Embodied activity is situated', it can be generalised that all TUI's are in fact situated given their tangible nature, and that in this respect both TUI's and EI's are similar.

The second criterion demands that embodied interaction is timely. By timely it is meant that real-world events require real-time interaction, so in the case of an interface, the interface would have to allow the user to respond in a timely manner, and not allow the flow of interaction to be broken. TUI's can differ from being very timely to not at all timely. An example of a timely system is the Illuminating Clay project [15] where you can change the input (the clay) and the output (the clay) immediately reflects the new state of the system. A non-timely example would be a TUI that didn't immediately update it's output in response to a change in the system or the users input.

The third criterion states that embodied interaction requires the use of multiple sensory modalities. As discussed in section 4.1, TUI's are in general multimodal, so in this respect they shall be regarded similar.

The fourth criterion is 'embodied activity is engaging'. In this respect, TUI's are different from EI's as they need not be engaging. It must be remembered that the concept 'engaging' not only considers the attention span of the user, and how occupied they are with the interface, but also by how much the interface needs the involvement of the user to function. A good example of a TUI that requires the user's ongoing attention is the topographic torch [2]. The topographic torch is an egocentric mapping device that displays the section of a map that refers to the area directly ahead of the user. This requires the user to move their body in alignment with where they are interested in.

From this comparison it can be seen that TUI's can differ from EI's in two respects, timeliness and engagement, whilst they are similar in the two respects of multimodality and situatedness.

5 Representing the Tangible-Enactive Space

It is proposed above that TUI's differ from EI's in two main aspects, timeliness and engagement. These are not simple binary states, but continua both from non-engaging through to fully engaging and from non-timely through to very timely. It is hence suggested here that it is possible to create a graph with 'engagement' for one axis and 'timeliness' for the second axis, on which it would be possible to place any TUI, with the result that it will be possible to plot how enactive a TUI is (fig. 1).

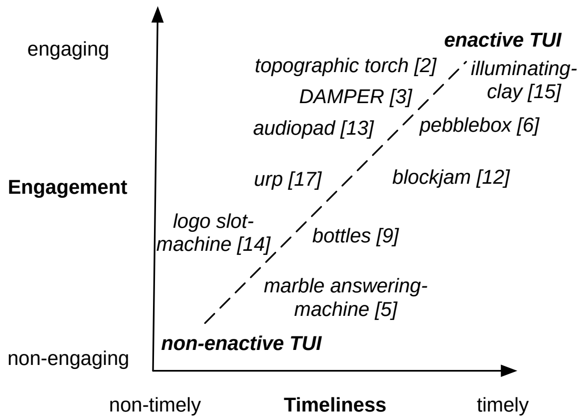


Figure 1: Graph showing the Tangible-Enactive Space, with project examples.

The utility of this graph is that designers of TUI's can gauge the enactive potential of the system they are designing, and modify their design so as to achieve the desired position within the tangible-enactive space. For instance the Logo slot machine [14] does not need to be fully enactive, as it is predominantly concerned with the symbolic manipulation of programming commands. Bishop's marble answering machine [5] is also placed in the non-enactive corner. The DAMPER [3] on the other hand, a tangible interface for music performance, benefits from being both timely and engaging and has hence been designed to enable enactive interaction. Similarly, the Pebble-Box [6] a musical interface designed for granular synthesis is placed in the enactive-TUI corner as it is both engaging and timely.

6 Discussion

When discussing the joining of TUI and EI theory, three questions come to the fore:

Does an enactive TUI make a better TUI? This is not necessarily the case. In some cases the ability for the transmission of enactive knowledge is desirable, however in some systems it is more desirable to create a purely iconic or symbolic system.

Are all TUI's enactive? The question of whether all TUI's are inherently enactive is answered to a certain extent by the Engagement vs Timeliness graph (Fig.1). The graph shows that only TUI's that are entirely unengaging or untimely are not enactive. Any TUI that is fully engaging and fully timely can be considered an enactive interface. So from this it is possible to see that only some TUI's make the status of being fully enactive (this can be seen as the 'strong' definition of an enactivity), and the majority are partially enactive ('weak' definition of enactivity), thus leaving only a small minority of TUI's that are truly un-enactive.

Are all enactive interfaces tangible? Although it is easy to jump to the conclusion that all EI's are tangible as in general a great deal are, it is possible to think of an EI that does not involve the manipulation of a tangible object. An example of this is a system that uses someone's free movement in space as input, for instance a typical virtual reality setup. The VR system allows the user to work in the enactive realm; although rather than utilising object manipulation skills it is possible to use a have a system where proprioceptive skills are more important.

7 Further Work

We are currently working on developing a series of musical interfaces that will be designed to explore the role of enactive interaction and how it can be built into TUI's. It is proposed that Fishkin's taxonomy be investigated with greater depth in regards to how metaphor and embodiment effect the enactive potential of these interfaces.

This discussion would be easier to carry out with a better understanding of what is, and what is not an enactive interface. It may be possible to get a better understanding of how to create an enactive interface by first trying to design a truly un-enactive TUI.

8 Conclusion

This paper has investigated the link between tangible user interfaces and enactive interfaces, showing how the two theories can be placed in a single design space. There is a lot more work to be done in this area and this paper will hopefully encourage further investigation into the link between the two disciplines.

References

[1] Armstrong, N. "An Enactive Approach to Digital Musical Instrument Design", PhD thesis, Princeton University, Nov. 2006.

[2] Bennett, M., O'Modhrain, S., and Quigley A. "Here or there is where? haptic egocentric interaction with topographic torch", CHI workshop proc., 2006.

[3] Bennett, P., Ward, N., O'Modhrain, S., Rebelo, P. "DAMPER: a platform for effortful interface development", NIME, NY, 2007

[4] Bruner, J. Processes of cognitive growth: Infancy, Worcester, MA:Clark University Press, 1968.

[5] Crampton Smith, G. "The hand that rocks the cradle" ID magazine May/June 1995, pp.60-65.

[6] Essl, G., and O'Modhrain, S. "An enactive approach to the design of new tangible musical instruments", Org. Sound, vol.11 pp.285-296, 2006.

[7] Fishkin, K.P. "A Taxonomy for and analysis of tangible interfaces", Pers. Ubi. Comp., 2004.

[8] Holmquist, L.E., Redström, J., and Ljungstrand, P. "Token-based access to digital information", lecture notes in Comp. Sci., 1999.

[9] Ishii, H., Mazalek, A., and Lee, J. "Bottles as a minimal interface to access digital information", CHI, pp.187-188, NY, 2001.

[10] McNerney, T.S. "Tangible programming bricks: An approach to making programming accessible to everyone", masters thesis, MIT media lab., 2000.

[11] Moggridge, B. Designing Interactions, MIT press, 2006.

[12] Newton-Dunn, H., Nakano, H., and Gibson, J. "Block jam: a tangible interface for interactive music", NIME, Montreal, 2003.

[13] Patten, J., Recht, B., and Ishii, H. "Audiopad: a tag based interface for musical performance", NIME, Dublin, 2002.

[14] Perlman, R., "Using computer technology to provide a creative learning environment for preschool children", MIT Logo memo # 24, MA, 1976.

[15] Piper, B., Ratti, C., and Ishii, H., "Illuminating clay: a tangible interface for landscape analysis", CHI, pp. 355-362, NY, 2002

[16] Ullmer, B., and Ishii, H., "Emerging frameworks for tangible user interfaces" IBM sys.jou., vol.39: Nos 3&4, 2000.

[17] Underkoffler, J., and Ishii, H. "Urp: a luminous-tangible workbench for urban planning and design", CHI, pp. 386-393, NY, 1999.

Spatial interaction in ambient communication

Stanislaw Borkowski Gilles Privat
Orange Labs/France Telecom R&D[1]
stan.borkowski@orange-ftgroup.com, gilles.privat@orange-ftgroup.com

Abstract

Ambient communication supports remote inter-personal communication by emulating social proximity-based protocols and arbitrates smoothly between face to face and remote communication. The position of users is taken into account together with other context elements supporting implicit context-aware interaction, using the distributed interaction and context acquisition resources available in a smart space. A new repertoire of spatial interaction functionalities affords implicit or explicit control with the communication system and provides contextual and control feedback integrated with the content of visual communication proper. Unshackling communication from device-based interfaces in this way amounts to a new form of enaction, supported by all interaction resources distributed in the smart space: enaction applies to interaction with the communication system in a way that emulates communication proper, letting ideally the communication system become transparent. This has the potential to open up new usage modes of communication for users to experiment with.

1. Beyond mobile and immersive communication

Ambient intelligence has been widely publicized as a sweeping paradigm shift, whereby embedded/ubiquitous ICT re-centers on users by becoming context-aware, personalized, adaptive and anticipatory [1]. What we call here *ambient communication* [8] projects the application of this broad vision to interpersonal communication.

Mobile communication has been extensively and brashly successful in making space *irrelevant*. In a different vein, immersive communication purports to detach users from their own physical space by immersing them in a simulated remote location that claims to be as close as possible to the original it represents.

Ambient communication is fundamentally different from both of these and takes a firm stance on *re-grounding* communication in the here-and-now familiar space and time of users, even though it draws upon the same technology infrastructure that supports mobile and immersive communication.

Mobile communication has freed users from wireline network attachments, but ties them to a handheld device that tends to morph into some fetishistic appendage or communication prosthesis… Immersive communication makes communication take precedence over vernacular activities, and, like virtual reality, forces the user to meet the information/communication world on its own sci-fi-inspired terms.

Ambient communication aims at freeing users from those ballyhooed devices and at bringing back communication firmly into our world, making it possible to communicate transparently, in a way similar to face to face communication, focusing on communication proper instead of focusing on devices that support communication. Communication services offered to users should no longer be exclusively linked to one of these devices and should no longer impose their own interface model.

Relieved from these constraints, ambient communication may become again an activity *situated* in the space that surrounds it, utilizing that most natural of human (and animal) capabilities: appropriating the surrounding space and making sense of it by moving our body around. The ensuing location/position/posture changes of people in this space become relevant as contextual features of interaction, in much the same way as they do in face to face communication. This situated communication is harmoniously integrated with other physically-grounded activities that take place in this space, over which it does not take precedence.

[1] Work partially supported by IST-FP6 Amigo project : *http://amigo-project.org*

2. Emulating proxemics in remote interaction

Most informal social situations involve an enactive rather than rule-based form of cognition, with tight feedback loops intertwining different sensorimotor channels between participating individuals [6]. We focus here on the spatial dimension of this communication, for which different proxemic patterns rooted in atavistic cultural backgrounds [4] condition both the human perception of space and the protocols of interaction. The sensorimotor loop involved between individuals comprises here perception of space/distance on the one hand and motor action of whole body movement or posture on the other hand. These proxemic constraints can be taken into account, and, to some extent, learnt, by monitoring reactions of other people when one gets more or less close to them in social interaction.

Let's illustrate things with the following scenario: Kate approaches Chen and Stan who are in face to face conversation. Subtle spatial cues (such as her relative distance to either of them, the orientation of her body, retreating slightly after first getting perhaps too close, etc.) may indicate whether she wants to interrupt them or not, whether she is interested in initiating communication with either one of them separately, or the two of them. Of course, other non-spatial cues and non-verbal communication channels come into play in such interaction. But the spatial-proxemic dimension is, among the channels opened up in this socially distributed form of cognition [5], the most salient and the most directly embodied [2], even when no direct contact is involved. The bodies of other people act as the external props or the spatial scaffolding that supports, grounds and bounds interaction occurring at other levels.

Now suppose Kate is in a different location from Chen and Stan. This kind of proximity-based trial and error on her part will not work with most communication setups, and remote communication will most commonly, as per a vestigial habit, take precedence over face to face communication in an abrupt and context-blind way.

Situated communication aims to offer the means for remote interaction to let this spatial dimension come into play in this natural way, similarly to what is possible in co-located interaction.

3. Context-aware communication supported by smart spaces: towards enactive context

Smart spaces provide the technological infrastructure to support the ambient/situated communication vision described above.

An assemblage of sensor& actuator devices available in the environment provides a backplane for supporting interactions that may occur throughout the environment. These devices may play a dual role: they handle direct explicit user interaction by either capturing input from users and rendering content to them in a traditional way and/or they may handle at the same time implicit/peripheral modes of user interaction as they capture and render physical context data that is relevant to user interaction. They are federated and managed jointly in both of these roles, as a unified distributed user interface, and as a set of context acquisition devices in an integrated context management system [9].

Context-aware communication complements both the *content* of communication proper (the audiovisual streams captured and rendered through the main interfaces of the communication system) and its *control* mechanisms.

The way context is often used in both cases draws heavily upon regular AI methodology, using a priori delineation and modelling of context, aggregation/abstraction of numeric data into logic variables and rule-based inference. Requiring an ad hoc set of rules and manual programming, this approach is clearly brittle and hardly scalable, especially when applied to high-level context that is intrinsically ambiguous.

Besides these practical problems, formalization of context is at variance with the proposed viewpoint that tends to recenter on *situation* in the physical world, at the expense of an objectivist and representational perspective coming from symbolist AI. This evolution conjoins the embodiment/enaction stance of cognitive sciences and the evolution of information and communication technologies towards a much stronger and richer coupling between the information world and the physical environment, human interaction being no longer confined to the desktop or other similar impoverished representations of the physical world. Environmental /physical context becomes prevalent in this view, and it is not coincidental that the dominant, if mundane and simplistic, use of context-awareness in telecommunications is in location-based services.

This post-cognitivist phenomenological context defines and gives meaning to the user's activity and at the same time, its basis and perimeter are continuously redefined by this activity [3]. Captured in an enactive loop, the context of communication *emerges* from communication/interaction, and provides the *grounding* of this interaction. Context could even be defined as all that is NOT reducible to rules or explicit modelling in communication, and is precisely what autistic individuals have trouble in capturing [6], making natural spontaneous communication all but impossible to them.

To some extent, the aim of context-aware communication could be to make the distinction between content, control and context irrelevant.

Primary content may be enriched at will with context comprising all kinds of ancillary data conveyed in face to face communication. Context is also used for implicit interaction, not necessarily in lieu of direct explicit interaction, but to complement it and blend with it. This cannot be achieved by pre-defining what kind of context is relevant and who to handle it.

Spatial interaction provides a way to show how this can be done, by integrating interaction and communication in a tightly coupled enactive loop.

4. Spatial interaction functionalities for visual communication

We are currently experimenting the use of low-level context information for interacting with an audio-visual communication setup supported by a smart space, with distributed context and content interfaces as described in the previous section The interpretation of higher-level context, e.g. social situation, is left to the users of our system. We draw on the concept of enaction to devise functionalities that are natural to use when interacting with the communication system, and that correspond to behaviors typical of face to face communication. We describe three such functionalities in the following subsections.

4.1 Follow-me

The idea of follow-me traces back to early projects [10][7], that proposed location-dependent access to interaction resources. In this case, a session opened by a user is moved to the closest computer. The user is free from carrying interaction devices and can interact with the system using any access point. We apply the idea of *re-grounding* the interaction in the physical space to interpersonal distant communication.

In video communication systems, cameras are used to capture the communication proper. However, cameras can also be considered as a potentially very rich source of context information. In our ambient communication prototype we use cameras to detect people in the proximity of camera-display sets. This information is managed by a context management system as described in [9]. When a user engaged in a conversation with a remote buddy changes location in the environment, the communication channel follows the movement by switching between available cameras and displays.

The follow-me feature allows the user to control a communication system composed of distributed interfaces in a very natural way. The reconfiguration of the system is transparent to the user, who only realizes that changing location in the environment does not hinder the engaged communication. Follow-me can also be applied in information-to-person communication services like video on demand.

4.2 Proximity-based zoom control

The coupling of the location with the control of the application can also influence interactions between distant users. In the Mirrorspace project [11], the changing distance of a person to a camera-display set has the same effect as a change of quality of the visual communication channel. As a person approaches, the displayed "reflection" of the user becomes focused and the initial blur disappears. In Mirrorspace the proximity is used as a means to control the amount of disclosed information between distant places. This coupling can be simplistically considered as an enactive way of turning the communication channel on and off.

Similarly to Mirrorspace, our prototype exploits proximity as context information controlling the transmitted data. The distance of a user engaged in a communication is linked to the control of the zoom to the distant location. When the user approaches the display showing a distant site, the corresponding view gets zoomed in. This view may get zoomed out (become wide-angle) again when the user moves back away from the display. In this way, the context information modifies not only the application, but also influences the interpersonal communication between users. Just like in collocated conversations the user can move from a "wide angle" discussion with multiple interlocutors to a "focused" private conversation with a privileged person.

This coupling also helps the user to better understand the context of the remote person. By simply moving back and forth, the user can adjust the view of the remote site to see also its periphery.

4.3 Trajectory-based interlocutor selection

The enactive interaction with an ambient communication system can be taken one step further when taking in account not only the proximity, but also the direction from which the user approaches communication devices. We explore this feature as a means for "selecting" an interlocutor from a remote group. In the scenario presented in section 2, Kate could indicate that she wants to talk to Stan rather than to Chen either statically by being closer to Stan or orienting her body towards him, or dynamically by adopting a trajectory towards him rather than towards Chen. This common social behavior can be, to some extent, supported by a remote communication system.

In our prototype, when a user approaches a display showing a group of remotely collocated people, the approach angle is translated to a selection of one of the remote person. The selected person is highlighted by a light boundary that fades away if the approach angle changes. In the remote site, the live image of the approaching user is shown together with a semi-transparent picture of the selected person. The transparency of the picture is eliminated as the selecting user approaches her/his display. The picture

is a very salient feature for the selected person, but the "called" person is free to leave the group and approach the "newcomer" to enter a private communication, or to ignore the remote person and continue chatting with the group.
The last of the presented functionalities requires real time video segmentation and target identification, which, in our current implementation, works only under controlled conditions. Similarly, the follow-me feature deployed in a multi-user environment requires user identification that we do not yet support. In "known" environments such as home or office, this could be implemented by a complementary analysis of visual context such as face or gait recognition.

5. Conclusions: spatial interaction as enaction

The spatial interaction functionalities presented above are, in a sense, the ultimate in *embodied* interaction; they dispense altogether with the intermediary *controls*, be they virtual or physical, that users manipulate in all traditional direct manipulation interface models: the user's *body* IS the control.

Enaction comes into play at two different levels: user to system interaction, and user to user interaction. Interaction with the interface system involves a close feedback between the motor action (whole body movement) and system output (perceived as visual content/context). A counterargument could be that the difference between this visual-motor feedback loop and the pointer/mouse feedback in the WIMP model is only a matter of degree and that moving one's body is not so different from moving a mouse[2]. Yet this person–to-system enaction is only an intermediary that both supports and evokes, as closely as possible, the person-to-person enaction involved in distributed social cognition [6].

This leads us to our main argument: contrary to the WIMP model, this interface model is *non-representational*, and it blends context and content as well as context and control in a way that cognitivist HCI cannot achieve, because context itself is not representational[3].

Controlling an interface by body movement makes implicit interaction partially explicit: when the user notices how the system reacts to actions that may at first have been non-intentional towards the system, she/he may divert these actions towards explicit interaction. Nevertheless, the system still uses relevant context that becomes blended with control.

In the "zoom by proximity" functionality, the user may also make context into content an vice versa: the larger panoramic image of the remote locale that she/he obtains by moving away from the interface resituates interaction in its larger context, after which she/he may again zoom in on what is most relevant in the scene.

Our ambient communication system supports social protocols by opening up the interaction repertoire, making it possible for users to attune themselves to those features of context that matter to them. They can play with the system and appropriate it in such a way that its use may become entirely transparent, as a perfect tool should be.

[2] A counter-counterargument (that we will not draw upon) could be that a few million years of evolution have made it more natural for primates to move their body around than to shift a nicely rounded object, be it called a mouse, on a flat horizontal surface…

[3] It is precisely because autistic spectrum disorder patients separate content from context that they cannot master naturalistic social cognition effectively [6]

References

[1] Aarts E. "Technological Issues in Ambient Intelligence", in Aarts & Marzano, (ed.)*The new everyday, Views on Ambient Intelligence"*, 010 Publishers, Rotterdam, 2003

[2] Anderson, M. L. "Embodied Cognition, a field guide", *Artificial Intelligence* 149(1):pp. 91-130

[3] Dourish, P." What we talk about when we talk about context", *Personal and Ubiquitous Computing*, vol. 8 n° 1, 2004

[4] Hall E. T. *The Hidden Dimension*, Doubleday, 1969

[5] Hollan, J, Hutchins, E. Kirsh D. "Distributed Cognition: Toward a new foundation for Human-Computer Interaction", *ACM Transactions on Human-Computer Interaction*, vol. 7 No 2, June 2000

[6] Klin A, Jones W., Schultz R., Volkmar F. "The enactive mind, from action to cognition: lessons from autism", *Phil. Trans. R. Soc. Lond.* B (2003) 345-360.

[7] Pingali, G. Pinhanez, T. Levas, R. Kjeldsen, M. Podalseck, "User-Following Displays", *Proc. of the IEEE International Conference on Multimedia and Expo 2002 (ICME'02)*, Lausanne, Switzerland, August 2002

[8] Privat G. "Ambient Communication: when devices disappear". Invited paper, *7th Workshop on Man-Machine Interfaces*, Berlin, October 2007

[9] Ramparany, F., Poortinga R., Stikic M., Jörg Schmalenströer J., Prante T., "An open Context Information Management Infrastructure", *Proceedings of the 3rd IET International Conference on Intelligent Environments* (IE'07), Ulm, Germany, September 2007

[10] Richardson T. "Teleporting - Mobile X Sessions", *In Proceedings Ninth Annual X Technical Conference*, Boston MA, January 1995

[11] Roussel, N., Evans H., and Hansen H. "Proximity as an interface for video communication." *IEEE Multimedia*, 11(3):12-16, July-September 2004

Tactile rendering of virtual objects

Alan C Brady Jianguo Qu
Sarah Carr Ian R Summers
Biomedical Physics Group, School of Physics, University of Exeter, UK
E-mail: i.r.summers@exeter.ac.uk

Abstract

Using the hands for active exploration of an object can provide information about the object's surface texture and about surface features such as edges and corners. In a virtual scenario, such information can be delivered by an array of contactors on the skin. Tactile rendering is a software description of surface properties which specifies appropriate drive signals for the array during active exploration of a virtual object. The design of such an array is discussed, together with a possible strategy for tactile rendering

1. Introduction

The tactile aspects of a virtual object may be represented as a spatial distribution of synthetic touch sensations over the fingertips. These sensations can provide information about the surface texture of the virtual object and about the contact between object and skin (contact area and position of edges/ corners).

To excite the skin mechanoreceptors, an array of contactors on the skin may be used to provide spatiotemporal patterns of mechanical input to the skin surface. Encounters with virtual objects, during active exploration of the workspace by the user, produce appropriate patterns of tactile stimulation on the fingertips.

When presenting the tactile aspects of a virtual object, the intention is not to reproduce the significant features of the small-scale surface topology of the object in terms of a virtual surface. Instead, the intention is to reproduce the perceptual consequences of small-scale features of the surface topology, i.e., appropriate excitation patterns over the various populations of touch receptors in the skin.

2. Design of a stimulator array

The optimal spacing of contactors in a stimulator array is determined by the spatial acuity of the sense of touch – around 1 mm on the fingertip [1]. However, a contactor spacing of 1 mm equates to around 100 contactors over the fingertip, each of which requires independent control. This is realistic for a passive (non-moving) device [2, 3] but is difficult to implement in an active device, for which a spacing of around 2 mm (i.e., around 25 contactors on the fingertip) may be a better choice. (There is some evidence [4] that perceptions available from an array with 2 mm spacing are not very different than those from an array with 1 mm spacing.)

In order to produce "realistic" touch sensations, a working bandwidth of around 10 to 500 Hz is required for the drive mechanism of each contactor, corresponding to the frequency range over which the various mechanoreceptors are sensitive [5]. Pacinian receptors are expected to respond most strongly to frequencies in the upper part of this frequency range (100 to 500 Hz, say); stimulation at lower frequencies is expected to stimulate mainly non-pacinian receptors. "Comfortable" sensation levels are produced by amplitudes of a few microns at frequencies around 300 Hz and a few tens of microns at frequencies around 50 Hz.

Design requirements for contactor spacing, working bandwidth and output amplitude may be satisfied by a variety of electromechanical drive mechanisms. Hafez and colleagues [6, 7] have developed arrays of drivers, based on shape-memory alloy or moving-coil technology, which apply normal forces to the skin. Hayward and colleagues [8, 9] have used piezoelectric-bimorph actuators to apply tangential forces. Summers et al. [10] have used similar actuators to apply normal forces, as have Kyung et al. [11].

The stimulator array developed in the ENACTIVE network and the HAPTEX project is shown in Figure 1. Piezoelectric bimorphs are used to drive 24 contactors in a 6 × 4 array on the fingertip, with a spacing of 2 mm between contactor centres. It can be seen that the drive mechanism is placed to the side of the finger and ahead of the finger, rather than below the contactor surface (which, at first sight, appears to

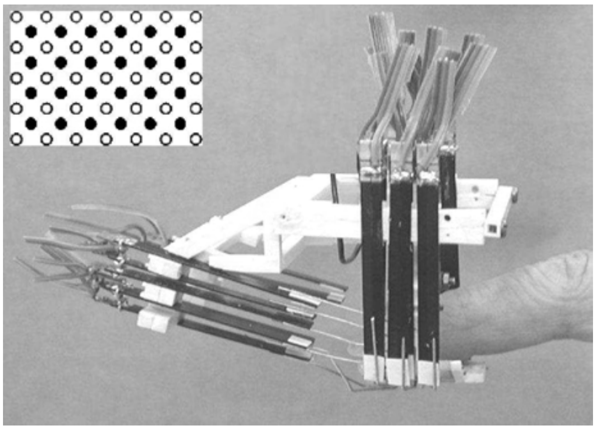


Figure 1. Stimulator array developed in the ENACTIVE network and the HAPTEX project. The contactor surface lies under the finger – contactors are driven by piezoelectric bimorphs (appearing as black rectangles). The inset shows the arrangement of 24 moving contactors, interspersed between the fixed contactors which transmit global forces to the fingertip.

be the most convenient location). With one such array on the index finger and another on the thumb, this positioning of the drive mechanism allows the finger to move close to the thumb so that a small virtual object can be manipulated between the tips of finger and thumb.

The contactor surface delivers to the fingertip the small forces associated with touch stimuli, but it must also deliver the larger forces associated with the overall mechanical properties of the virtual object, represented by the output of a force-feedback system. However, the moving contactors which provide touch stimuli are driven by delicate piezoelectric mechanisms and so they are unsuitable for delivering the force-feedback output, which may involve forces of considerable magnitude. Consequently, the contactor surface includes an additional set of contactors ("fixed" contactors – see inset to Figure 1) which deliver the force-feedback output, in parallel with the tactile stimulation from the moving contactors.

3. Drive signals for a stimulator array

During active exploration of a virtual tactile environment it is necessary to generate in real time a drive waveform for each contactor of the stimulator array(s) which are in contact with the user's fingertip(s). The amount of data which must be generated "on the fly" is thus considerable. For example, the ENACTIVE/HAPTEX system has 24-contactor arrays on finger and thumb, requiring 48 analogue drive signals, in principle each with a bandwidth of around 500 Hz. However, because of the limited temporal resolution, frequency resolution and phase sensitivity of human touch perception [12, 13, 14, 15, 16], there are possibilities for a significant reduction in the data flow. For example, each drive signal may be reduced to the sum of a limited number of sinusoidal components, distributed across the working bandwidth (10 to 500 Hz – see above). The drive signal may then be simply specified in terms of the amplitudes of these components, which require an update every 20 ms or so.

In the ENACTIVE network and the HAPTEX project, a cut-down version of this scheme has been developed, in which the drive signal to each contactor is the sum of components at only two frequencies: 40 Hz and 320 Hz. Following the suggestion of Bernstein [17], the higher frequency was selected (at 320 Hz) to target pacinian receptors and the lower frequency was selected (at 40 Hz) to target non-pacinian receptors. Each drive signal is specified by the sinewave amplitudes A_{40} and A_{320} of the two signal components. These are updated every 25 ms (once per cycle for 40 Hz, once per 8 cycles for 320 Hz). A virtual tactile surface is specified in terms of an amplitude map for each of the two frequency components that make up the stimulus.

4. Tactile rendering

During exploration of a virtual tactile environment a drive waveform is specified for each contactor of the stimulator array(s) – see above. A significant problem is the current lack of knowledge on the origin and nature of excitation patterns in real situations of tactile exploration of an object. The mechanical stimulation of a given receptor has a complicated relation to the mechanical properties and topology of the object's surface, to the mechanical properties of the skin and its local topology (especially skin ridges, i.e., fingerprints), and to the precise nature of the exploratory movement (speed, contact pressure and direction). Although it may be possible to produce an accurate software model of an object's surface, it is not at present possible to augment this with an accurate model of the skin/surface interaction. This situation may change in the near future: research is currently underway to develop an "artificial finger" with embedded transducers to mimic mechanoreceptors; improved finite-element models may also provide useful data.

For the particular case of the manipulation of textiles, the situation is more promising: Information on the nature of the mechanical input to the skin's mechanoreceptors is available from the Kawabata system for evaluation of textiles [18]. This provides a range of data on the textile sample under test, including surface roughness and surface friction profiles which are direct measures of the mechanical excitations produced when a probe is moved over the textile surface. The probe and associated instrumentation are designed so that the measured quantities correlate well with subjective assessment of the textile surface. Hence the Kawabata surface measurements provide an approximation to the “perceived surface”, i.e., the surface after it has been “filtered” through the surface/skin interface. They thus provide a good basis for specifying drive signals for a stimulator array, in order to provide the tactile component for a virtual textile. Kawabata measurements have been used in this way by Govindaraj et al. [19]; they have also been used to provide source data for the tactile rendering developed within the HAPTEX project on virtual textiles. Preliminary work on tactile rendering within the HAPTEX project is described by Allerkamp et al. [20]. More recent ideas are presented below.

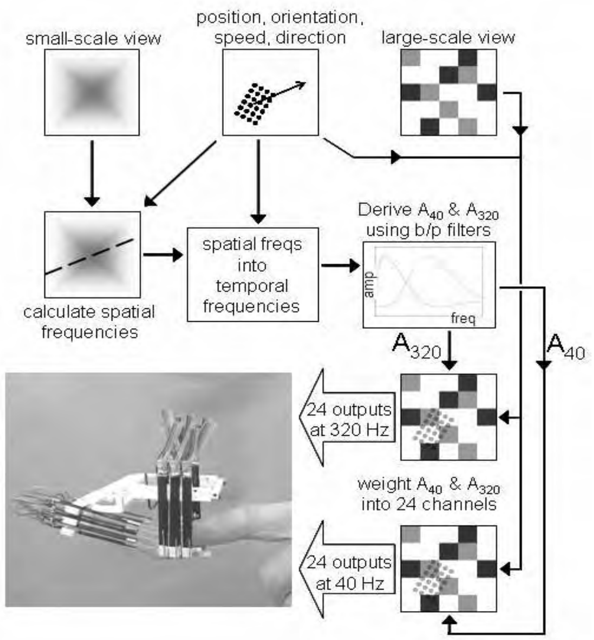


Figure 2. The scheme for tactile rendering which has been developed in the HAPTEX project

Figure 2 outlines the scheme for tactile rendering, by which the drive signal to each point in the stimulator array is specified in terms of amplitude A_{40} at 40 Hz and an amplitude A_{320} at 320 Hz (see above), and these amplitudes are in turn specified by the interaction between the virtual object and the exploratory movements of the user. For each digit, the tactile renderer generates 24 drive signals for the 24 contactors of the stimulator array. Input and output are specified in 25 ms timesteps. The input data are:

♦ a small-scale description of the object surface, represented as 2D *k*-space, derived from a pseudo-topology at around 0.01 mm resolution over an area of a few mm^2;

♦ a large-scale description of the object surface: a representation of the non-uniformity of the surface, specified as pseudo-amplitudes at 1 mm resolution over an area of several tens of cm^2;

♦ position and orientation of the finger pad on the virtual surface;

♦ speed and direction of the movement of the finger pad over the virtual surface.

The operation of the renderer is as follows: Taking account of the direction of movement, a spatial-frequency spectrum is calculated from the 2D *k*-space of the small-scale description of the virtual surface. Information about the speed of movement of the finger pad is used to convert spatial-frequency components into temporal-frequency components. The resulting temporal-frequency spectrum is reduced to only two amplitudes, A_{40} and A_{320}, by application of appropriate bandpass filter functions, corresponding to the 40-Hz and 320-Hz channels. (It should be noted that the signal-processing operations to this point may be performed only once per 25-ms timestep, i.e., they may be common to all 24 output channels.) Amplitudes for the 40-Hz component in the drive signals for each of the 24 channels are obtained from A_{40} by weighting according to data from the large-scale description of the virtual surface, for the 24 locations on the finger at which the contactors of the tactile stimulator are positioned. Similarly, amplitudes for the 320-Hz component in the drive signals for each of the 24 channels are obtained from A_{320} by weighting according to data from the large-scale description of the virtual surface. (Note that, in principle, different large-scale descriptions of the virtual surface may be used in the 40-Hz and 320-Hz channels, to allow for the observed difference in spatial resolution on the fingertip at the two frequencies.)

5. Discussion

When using stimulator arrays and rendering schemes as described above, the intention is to present time-varying spatial patterns of tactile stimuli which have two perceptual dimensions: one relating to intensity and one relating to spectral distribution. In order to establish the potential for such a system, it is necessary to determine whether a two-dimensional perceptual space can indeed be created in this way – it is very likely that the intensity dimension is available to the user, but less obvious that the spectral dimension is available. However, recent results from Kyung et al. [21] demonstrate that test subjects can detect changes of frequency when stimuli are presented via a stimulator array in an active task, so it seems that spectral information is indeed available in such a scenario.

Initial evaluations of the ENACTIVE/HAPTEX system (Figure 2) also suggest that a 2D perceptual space can be achieved. For uniform stimuli (i.e., stimuli with no spatial variation over the skin), the spectral dimension appears relatively weak – changes in spectral balance at constant subjective intensity tend to be less noticeable than changes in subjective intensity at constant spectral balance. (There are perhaps 4 to 5 discriminable steps of spectral balance along an equal-intensity contour.)

Perhaps the most interesting observation when using the ENACTIVE/HAPTEX system is a strong interaction between the perceived spatial aspects of the texture and the stimulation frequency. If the stimulation frequency is changed from 40 Hz to 320 Hz, the perceived sensation during active exploration changes much more if the texture is spatially non-uniform than if it is spatially uniform. It is clear that the spectral dimension provides a significant enhancement to the available range of tactile sensations. Experiments are currently under way to investigate the perceptual space in detail, and to further investigate the interaction between the spatial aspects of the texture and the nature of the perceptual space.

Using physical data from a selection of real fabrics (obtained with the Kawabata system), the ENACTIVE/HAPTEX system has been used to simulate the tactile aspects of those fabrics. Given the apparent mismatch between the real situation (fingertip touching a textile) and the virtual situation (fingertip touching the metallic contactors of a stimulator array), results are surprisingly good – in some cases test subjects are able to match real and virtual textiles in terms of their tactile qualities.

Acknowledgements

This work is funded under the Sixth Framework Programme (FP6) of the European Union, as part of the ENACTIVE Network of Excellence (Contract No. FP6-002114) and the HAPTEX project on virtual textiles (Contract No. IST-6549).

References

[1] Johnson K.O., Yoshioke T., Vega-Bermudez F. "Tactile functions of mechanoreceptive afferents innervating the hand", J. Clin. Neurophysiol. 17, 539–558 (2000).

[2] Summers I.R., Chanter C.M. "A broadband tactile array on the fingertip", J. Acoust. Soc. Am. 112, 2118–2126 (2002).

[3] Killebrew J.H., Bensmaia S.J., Dammann J.F., Denchev P., Hsiao S.S., Craig J.C., Johnson K.O. "A dense array stimulator to generate arbitrary spatio-temporal tactile stimuli", J. Neurosci. Meth. 161, 62–74 (2007).

[4] Summers I.R., Chanter C.M., Southall A.L., Brady A.C. "Results from a Tactile Array on the Fingertip", Proc. Eurohaptics '01, Birmingham, 26–28 (2001).

[5] Gescheider G.A., Bolanowski S.J., Hardick K.R. "The frequency selectivity of information-processing channels in the tactile sensory system", Somat. Mot. Res. 18, 191–201 (2001).

[6] Benali-Khoudja M., Hafez M., Alexandre J.M., Kheddar A. "Electromagnetically Driven High-Density Tactile Interface Based on a Multi-Layer Approach", International Symposium on Micro-mechatronics and Human Science (MHS 2003), 147–152 (2003).

[7] Hafez M., Benali-Khoudja M. "3D Tactile Rendering Based on Bi (Multi) stable SMA Monolithic Systems", International Symposium on Micro-mechatronics and Human Science (MHS 2004), 93–98 (2004).

[8] Pasquero J., Hayward V. "STReSS: A Practical Tactile Display System with One Millimeter Spatial Resolution and 700 Hz Refresh Rate", Proc. Eurohaptics '03, Dublin, 94–110 (2003).

[9] Levesque V., Pasquero J., Hayward V. "Braille Display by Lateral Skin Deformation with the STReSS2", Tactile Transducer. Proc. World Haptics '07, Tsukuba, 115–120 (2007).

[10] Summers I.R., Brady A.C., Syed M., Chanter C.M. "Design of Array Stimulators for Synthetic Tactile Sensations", Proc. World Haptics'05, Pisa, 586–587 (2005).

[11] Kyung K.U., Ahn M., Kwon D.S., Srinivasan M.A. "A compact planar distributed tactile display and effects of frequency on texture judgment", Adv. Robot. 20, 563–580 (2006).

[12] Sherrick C.E. "A scale rate for tactual vibration", J. Acoust. Soc. Am. 78, 78–83 (1985).

[13] Rabinowitz W.M., Houtsma, A.J.M., Durlach, N.I., Delhorne, L.A. "Multi-dimensional tactile displays: Identification of vibratory intensity, frequency and contactor area", J. Acoust. Soc. Am 82, 1243–1252 (1987).

[14] Van Doren C.L., Gescheider G.A., Verrillo R.T. "Vibrotactile temporal gap detection as a function of age", J. Acoust. Soc. Am. 87, 2201–2205 (1990).

[15] Formby, C., Morgan, L.N., Forrest, T.G., Raney, J.J. "The role of frequency selectivity in measures of auditory and vibrotactile temporal resolution", J. Acoust. Soc. Am. 91, 293–305 (1992).

[16] Summers I.R., Whybrow J.J., Milnes P., Brown B.H., Stevens J.C. "Tactile perception; comparison of two stimulation sites", J. Acoust. Soc. Am. 118, 2527–2534 (2005).

[17] Bernstein L.E., Eberhardt S.P., Demorest M.E. "Single-channel vibrotactile supplements to visual perception of intonation and stress", J. Acoust. Soc. Am. 85, 397–405 (1989).

[18] Kawabata S., "The standardization and analysis of hand evaluation", 2nd ed. The hand evaluation and standardization committee, The Textile Machinery Society of Japan, Osaka, (1980).

[19] Govindaraj M., Garg A., Raheja A., Huang G., Metaxas D. "Haptic simulation of fabric hand", Proc. Eurohaptics '03, Dublin, 253–260 (2003).

[20] Allerkamp D., Bottcher G., Wolter F.-E., Brady A.C., Qu J., Summers I.R. "A vibrotactile approach to tactile rendering" Vis. Comput. 23, 97–108 (2007)

[21] Kyung K.U., Kwon D.S. "Perceived roughness and correlation with frequency and amplitude of vibrotactile frequency", Proc. Eurohaptics '06, Paris, 277–282 (2006).

A Real-Time Simulator for Virtual Reality conceived around Haptic Hard Constraints

CASTET Julien* COUROUSSE Damien* FLORENS Jean-Loup ‡
LUCIANI Annie*
(*)Laboratory ICA – INP Grenoble, France
(‡)ACROE – INP Grenoble, France
*E-mail: ** julien.castet@imag.fr, damien.courousse@imag.fr, jean-loup.florens@imag.fr, annie.luciani@imag.fr ***

Abstract

This paper presents a real-time physically based platform for multi-sensory interactive simulation. This platform is centered on high quality dynamic requirements driven by the concept of instrumental interaction. It is oriented towards the simulation of vis-à-vis human-object interactive simulation for a broad range of physical phenomena, with a specific focus on simulations with demanding capacities regarding dynamics such as tool use, object manipulation, music instrument playing, etc.

The platform consists in a precisely synchronized multiprocessor architecture extended with a DSP board. Two versions have been implemented corresponding to specific simulation requirements: (1) a highly reactive simulator for models with strong dynamical requirements but low computation needs, (2) a 64 bit floating point multiprocessor architecture for large 3D models with complex interactions such as fluids, smoke, crowds.

1. Introduction: classical instrumental VR

The introduction of haptic devices in VR platforms oriented towards manipulation has motivated lots of active researches the last 20 years. Like the geometrical modeling and light modeling, physically based modeling stage is also added in the simulation like an independent part.

For the classical instrumental VR from "Computer Graphics", technical problems are generally to generate force at least of 1 kHz and compute complex cinematic treatments at the same time in a coherent way. This part presents some examples of these platforms categorized in function of their computation load repartitions. Currently, it exists three types of VR platform architecture.

SPORE (Simulation of Physically based Objects for Real-time Environments) is an example of "centralized approach" for surgical simulation [1]. The simulator is composed of three units: a mechanical unit, a visual unit and a collision unit that run on one general-purpose PC with a Phantom haptic device. A minimal kernel is dedicated to common processes: ODE and collision detection. A part of the kernel a collection of complex physically based models is proposed.

SPRING is a surgical simulation system dedicated mainly for Collaborative task [2] based on the "Client-Server" configuration. The simulation process runs on a single computer, haptic and audio devices are connected to the simulator through Ethernet network and visualization is duplicated on different displays.

Finally there is the "distributed approach" platform where simulation runs on PC cluster. SIMNET [3] was one of the first platforms using the distributed approach; it was a multi-users platform for the training of shooting. We can also cite FlowVR, a development library recently used, providing tools for the development of interactive applications on clusters of general-purpose PCs [4].

These are the main architectural tendencies of general VR platforms. An important part of works VR are focused on interactive simulation, that is, to give means to the human user to interact with the virtual environment that is simulated. It exists different strategies of implementation. The next part introduces them and their main requirements and the last part describes our Real-Time Simulator for Virtual Reality.

2. Ergotic interaction for Virtual Reality

2.1. Gesture controllers and the mapping strategy

Very soon the computer appeared as a mean to make the human enter into new virtual worlds. The emergence of new human-computer interaction interfaces was very soon foreseen [5]. The first gesture interfaces were "manual-input devices", that is, motion sensors: mice, tablets, etc. This approach is still used a lot currently in the fields of HCI and computer music. One can summarize it as following: in mapping strategies (Fig.1), motion sensors feed the simulation process with position information from human motion

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19^{th}-22^{nd}, 2007

[13]. The simulation process gets position data or event-based information from the motion sensors. Gesture data flow from human to computer. This implementation strategy is mostly found into computer systems for electronic music, where the need is to provide an easily extensible set of very versatile gesture possibilities for the player. It is however criticized when there is a need for the human user to be in contact with the physical mechanism represented in the simulation of the virtual object [13].

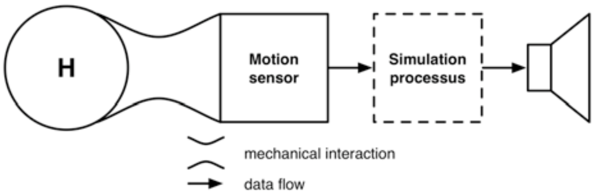


Figure 1. The mapping strategy

2.2. The haptic component in VR

Human interaction with an object or an environment necessarily involves mechanical coupling and exchange of mechanical energy. C. Cadoz first coined this particular property of human gesture under the term 'ergotic' [7]. In the case of computer mediation this situation involves the use of bidirectional transducers to connect the digital world of the simulation process to the physical world of sensory-motor phenomena. The existence of a bidirectional flow of information is a necessary condition for the implementation of 'ergoticity' in human-computer interface.

HCI science has shown that performance of human-computer interaction could be greatly improved by the use of haptic interfaces [8]. In the field of Computer Graphics, the haptic device is considered as a device capable to "render" shapes or rigidity of the objects simulated thanks to force feedback [9].

A different approach, which we support in our work, merely considers the haptic device as a mean to recover, in artificial interaction situations that are mediated by computer, the ergotic function that exists in natural human-object interaction. Within this approach several positions can be found, corresponding to different modeling paradigms of the instrumental chain.

In most of VR applications, the gesture interface becomes a haptic interface instead of a motion sensor. It is commonly admitted that the flow of data between the haptic device and the simulation process must be exchanged at sampling rates between 1 and 3 kHz. This simulation process includes the simulation of the mechanical part of the virtual object or environment that is directly related to the haptic device and to human gesture (Fig.2). Due to computation limitations and/or modeling choices, only a part of the simulation process is running at such frequency, and the rest of the simulation process is either running at lower frequency, asynchronously of the mechanical model, or is based on the triggering of pre-recorded samples controlled by events (generally for sound generation, as in the mapping strategy).

This implementation strategy is interesting in a large panel of situations because it leads to an efficient use of computation resources, and to suitable, yet not satisfying, solutions for the user, from the point of view of gesture interaction.

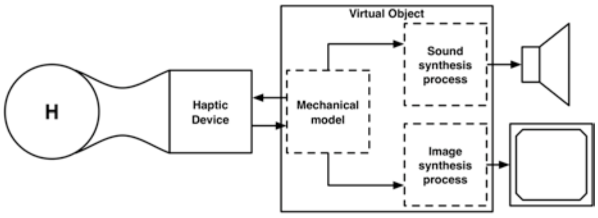


Figure 2. Common haptic implementation into VR

2.3. Instrumental VR

The common implementation strategy detailed above is not satisfying from the point of view of instrumental playing, as the energetic chain between the human and the simulated object, taken as an instrument, is cut in the simulation process [13]. Indeed, such implementation strategy can be satisfying when the mechanical part of the object that is in direct interaction with the human body is decoupled from the rest of the object. For example, in the mechanism of the piano, one can consider that the key mechanism is not structurally coupled to the vibrating structure of the piano (the strings and the soundboard). But there are situations where such decoupling is not possible without breaking the targeted phenomenon that is emerging from human-object mechanical coupling. This is the exemplary case of violin playing, or of having your wet finger making 'sing' a crystal glass. To simulate properly such situation, one cannot avoid having only one simulation process as the sounding part and the mechanical part of the object are intimately coupled (Fig.3). This simulation process is responsible for the generation of all sensory information involved in the simulation.

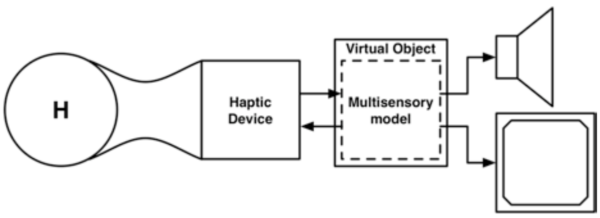


Figure 3. Our implementation of instrumental VR

2.4. System constraints for instrumental VR

In instrumental VR, the relation between the gesture interface and the simulated object is not based on phenomenological information, but rather on a bidirectional exchange of data flows that are synchronized with the simulation process (Fig3).

Dynamics of the physical phenomena simulated should be correctly represented into computer simulation, both in terms of simulation frequency and temporal latencies. Therefore, technical requirements are introduced so that instrumental playing is respected:

- The bandwidth of the simulation should encompass the cut-off frequency of the physical phenomenon that is simulated: if the simulated model includes acoustical parts, simulation bandwidth should be high enough to generate acoustical frequencies of the sound signal (10 to 50 kHz).
- Dynamic range of the physical variables should encompass the dynamic range of the reference physical phenomenon.
- The simulation process should be synchronized with devices involved in the interactive simulation, such as transducers (haptic device or loudspeakers). Especially, I/O latency of the simulation process should not exceed one simulation period as latency introduces physical distortion [10]. Time determinism is the only way to guaranty this synchronization: a step of simulation must be computed within fixed time windows.

3. Architecture of the platform

3.1. General presentation

This part presents the main features of the platform developed to satisfy presented constrains of ergotic tasks. Concerning the modeling framework, the ACROE team has designed since 1984 computer formalism, called the CORDIS ANIMA system [11], based on discrete mass-interaction modeling. A physical object or a set of physical objects are modeled and simulated as a network where the nodes are the smallest modules representing inertia (the <MAT> elements) and where the links (the <LIA> elements) represent physical interactions between them. The modules are all implemented with explicit algorithms, allowing for deterministic computation. Thus our multi-sensory simulation is based on one model composed of a large number of simple algorithms allowing a regular computation synonymous of determinism.

The input/computation/output sequence can also be easily synchronized on an external clock. Simulation frequency can be adjusted between 1 and 50 kHz according to the bandwidth of the physical phenomenon targeted.

Depending on the expected task and circumstances of platform use, simulation requirements are variable. This platform is also conceived like a modular hardware platform composed with different computation units.

One of the hardware components is multi-processor computers (bi or quadri) from Concurrent Computer Corporation. Processors used are AMD Opteron 2 GHz with cache of 1Go and 64 bits architecture.

A DSP board from Innovative Integration, called TORO board, is the second main component of the platform. The DSP embedded is the TMS320C6711 characterized by a computation frequency of 150MHz. This card provides 16 simultaneous analog inputs and outputs up to 250 KHz each, both at 16-bit resolution for high quality haptics. A/D and D/A converters are synchronized on the same clock signal as the simulation process and are used for the exchange of data with the haptic device (ERGOS panoply [22]).

Considering that commercial sound boards present non negligible latencies, we have also chosen to take benefit of the 16 bits precision D/A converters for sound outputs (up to 4 channels, for quadraphonic sound), allowing for very short latencies (less than 5 µs for a simulation frequency of 44,1 kHz).

These hardware components could be used together or independently providing a range of configurations for various performances. Two of them have been realized and described in the two next parts. A third configuration, presented in the last part, is currently developed.

3.2. High Reactivity simulation

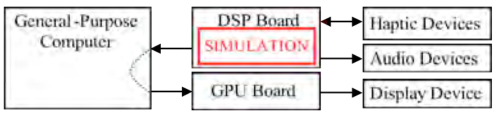


Figure 4. Configuration with DSP Board

This configuration is only based on the DSP-board with high channel of analog I/O. This module is particularly adapted to be used for models requiring hard reactivity and a low computation power. The simulation runs on the DSP and its host, a general-purpose computer, is only used for simulation control and for visualization. The emblematic example of this configuration is simulation of violin playing.

3.3. Simulation of complex 3D scenes

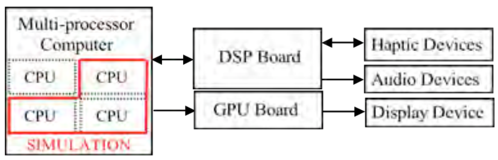


Figure 5. Configuration with multi-processor computer

Running the simulation on the multiprocessor system increases the computation power and so the complexity of scene simulated. Moreover our physical models simulation algorithms with its network topology are very pertinent for a multi-processor repartition. Indeed, the communication inter-processor is resolved with a specific physical module satisfying time and physical coherence called «ghost module». It consists in a semi-mirror to allow for network of models cutting (fig.6).

The combination of this Multi-processor architecture and the Operating System specialized for Real-time applications [12] allow satisfying time determinism requirements. Some low-level tools provide the means of controlling computation:

- Management of access right (memory, processor...)
- Processor shielding against process, interrupts.
- The deactivation of scheduler
- Assigning process to processors
- Protecting process execution against memory swap
- Inter-process synchronization (spin lock, active waiting…)

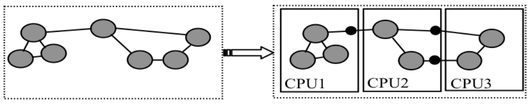


Figure 6. Repartition with «Ghost Modules»

A first evaluation of the system has been made with agglomerate, 10^3 masses and so 10^6 interactions. It can simulate this model with frequency of 500 Hz.

3.4. Simulation of complex scenes with high reactivity of haptic modality.

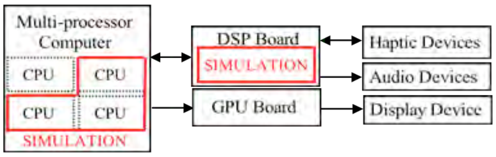


Figure 7. Configuration with multi-processor computer

The last solution consists in using every computation units: the Multiprocessor computer and the DSP board. Thus the distribution of model can be done in order to simulate a complex scene with high reactivity for haptic modality. For example, a large model can be simulated with 3kHz frequency on computer and connected to a cinematic transform simulated with 44KHz on DSP.

4. Conclusion

Unlike other Virtual Reality platforms, this new real-time physically based hardware platform allows a complete synchronous multi-sensory interactive simulation. Its hardware modular architecture provides a range of configuration adapted to different issues.

According to performances obtained with our different implementations on individual configurations (agglomerate, friction of bow, pebble box, smoke…), the connection between the multiprocessor computer configuration and the DSP board configuration is a really promising solution for a future virtual reality platform allowing the simulation of complex scene with high quality of haptic interaction.

Actually, this connection is in development and the optimization of PCI transfer between host and the DSP Board is the real last problem for our multi-frequency configuration.

References

[1] Meseure, P., and al (2003) “A Physically-Based Virtual Environment dedicated to Surgical Simulation”, International Symposium on Surgery Simulation and Soft Tissue Modeling (IS4TM), Sophia-Antipolis.

[2] Montgomery, K., and al (2002) “Spring: A general Framework for Collaborative, Real-time Surgical Simulation”, Medicine Meets Virtual Reality, IOS Press, Amsterdam.

[3] Calvin, J., and al (1993) “The simnet virtual world architecture”, IEEE Virtual Reality Annual International Symposium, pp. 450-455, Seattle.

[4] Dequidt, J., and al (2005) “Collaborative interactive physical simulation”, 3rd International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia Proceedings, Dunedin, New Zealand.

[5] Sutherland, I. (1965). The ultimate display. In Kalendich, W. A., editor, Proceedings of IFIP Congress 65, volume 2, pp 506–508, New York City. Spartan Books, Washington, D.C.

[6] Wessel, D. and Wright, M. (2002) “Problems and prospects for intimate musical control of computers”. Computer Music Journal, 26(3), pp. 11–22.

[7] Cadoz, C. (1994) “Le geste, canal de communication homme/machine. la communication instrumentale“. Technique et Science de l’Information, 13(1), pp. 31–61.

[8] Rosenberg, L. B. and Brave, S. (1996) “Using force feedback to enhance human performance in graphical user interfaces.” In Proceedings of CHI’96: Conference on human Factors in Computing Systems, Vancouver, Canada.

[9] Basdogan, C. and Srinivasan (2002) M. A., Handbook of Virtual Environments, pp. 117–134. Lawrence Erlbaum, Inc., London.

[10] Castagne, N. and Florens, J.-L. and Luciani, A. (2005) “Computer Platforms for Hard-Real Time and High Quality Ergotic Multisensory Systems - requirements, theoretical overview, benches”, Proceedings of ENACTIVE05 - 2nd International Conference on Enactive Interfaces, Genoa, Italy.

[11] Florens, J.-L. and Urma, D. (2006). Dynamical issues at the low level of human / virtual object interaction. In DBL, pp. 47, Arlington, VA, USA.

[12] Florens, J-L., and al (2004) “ ERGOS : Multi-degrees of Freedom and Versatile Force-Feedback Panoply”, pp. 356-360, Munchen, Germany.

[13] Florens, J.-L. and Cadoz, C. (1991) Representation of Musical Signals, chapter The physical Model, Modelisation and Simulation Systems of the Instrumental Universe, pp 227–268. The MIT Press, Cambridge, Massachusetts, USA.

Camera-based Gaze Control for Virtual Characters

Parag Chaudhuri George Papagiannakis Nadia Magnenat-Thalmann
MIRALab, University of Geneva, Switzerland
E-mail: {parag, papagiannakis, thalmann} @miralab.unige.ch

Abstract

Virtual characters have become crucial to many interactive 3D graphical virtual environment simulations in the context of virtual reality and computer games. One of the primary concerns of the designers of such environments is to model the interaction of the virtual characters with the users. The gaze of the virtual character forms a chief component of this interaction. When the gaze of the virtual character is directed correctly towards the users, they perceive a sense of presence inside the virtual environment, during their interactions with the virtual character. We present, in this paper, a simple idea to control the gaze of a virtual character based on the viewpoint of the user involved in the interaction, suitable both for Virtual Reality as well as Augmented Reality applications. We also place the idea in the context of methods to control gaze that already exist in literature by providing a summary of the existing works.

1. Introduction

Virtual environments have become ubiquitous in many applications of modern day computer graphics simulations. In particular, they are central to computer games and virtual reality simulations. The primary emphasis of all these simulations, which involve interactions with the user, is to immerse the user in the environment. A very direct and common way to achieve this is to let the user interact with virtual characters in the simulation rather than with inanimate objects.

The virtual character, must however, appear realistic to the user in action, looks and behaviour. A lot of work has been done on the visual plausibility of characters in recent times. However, once the user is allowed to interact with the virtual characters, the illusion of believability can disappear rapidly. Often the behaviour of the characters can look “wooden”, exhibiting perhaps, a lack of variety in response, a lack of emotional range or a lack of adaptation to the users’ attitude towards the character.

The subject of this paper is one such characteristic of the virtual character, namely the gaze, which goes a long way in establishing a sense of presence for the user in the virtual environment when interacting with the virtual character.

In human interactions, gaze serves a variety of functions [1][7]. It is used to gather information. It is used to regulate the information gathered from the environment. It is also used to regulate conversations. Gaze also serves to signal interest and express emotional state. This paper presents a simple way to model direction of gaze of a virtual character, during interactions with a real human.

We begin by presenting the existing work in gaze modeling for virtual characters in the next section. Section 3 we present our technique for computing the gaze direction. We conclude in Section 4 with a brief discussion and approaches for future work.

2. Background

Mixing virtual and real humans in a common Augmented Reality (AR) space based on narrative, dramaturgical interaction has been provided in [4] capitalizing real-time human-scale augmentations [8] for enhanced presence in these spaces. However, these examples did not allow for *mutual persistent presence* between real and virtual humans, since the virtual ones were not capable of sensing the real ones, looking at them and thus establishing a more advanced conscious relationship. The impact of eye gaze model on communication using humanoid avatars has been recently studied [6][11] based on different eye-head tracking sensors. In this work we aim to provide a simple and efficient model without the need of extra sensors, since we derive the basic camera matrix from the AR vision-based camera tracking inside-out algorithm [8]. Another recent approach for eye control was provided by [2] where the gaze model was considered as a subtask of neck control. Wearable eye trackers could also be considered for Virtual Reality (VR) applications [9][3] but would be impractical for

AR since the user already is equipped with a video or optical see-through HMD and an extra wearable tracker would render the mobile experience less practical.

3. Camera based Gaze Control

We propose a simple model to compute the gaze direction for a virtual character. We assume that a single user is interacting with the character. A camera mounted on a Head Mounted Display (HMD) of the user is used to record views of the real world. The real world is then augmented with the virtual character and the image is projected back onto the HMD. So the user can see the virtual character in front of their eyes (see Figure 1).

We track feature points video recorded by the camera in real-time. Feature points computed in one frame of the video are matched with feature points in subsequent frames. A stratified metric reconstruction from the computed correspondences gives the camera matrix, **P**, for the camera mounted on the HMD[5][10]. The 3×4 projective camera matrix, **P**, is decomposable as **K**[**R**|**t**], where **K** is a 3×3 matrix containing the focal length of the camera, **R** is a 3×3 submatrix controlling the view direction, and **t** is a 3×1 submatrix governing the viewpoint distance.

Figure 1. The user with a camera attached to the HMD

We recover the view direction v as:

$$v = det(\mathbf{M})\mathbf{m}^3,$$

M is the first 3×3 submatrix of **P** and $\mathbf{m}^{3T}$ is the third row of **M**. *det*(**M**) is the determinant of **M**. The camera center **C** can be estimated as the right null space of **P** by solving **PC** = 0. The look-at point, l, is given by

$$l = \mathbf{C} + \lambda v.$$

The view-direction, *v*, is constantly updated in real time as head of the user moves in order to look at the virtual scene. This view-direction is used to control the gaze of the virtual character interacting with the user as explained in Algorithm 1.

We assume that the scene is represented as a scene graph and the character is a part of this scene graph. The algorithm gives the method we use to compute the gaze direction of the character during the rendering of the scene.

Algorithm 1: *Algorithm to gaze of the virtual character.*

Require: Scene graph **S** containing the virtual character and view-direction *v*.

1. **while** **S** is being rendered on the HMD
2. **do**
3. Create a node visitor, $\mathbf{n}^{\mathbf{v}}_{\mathbf{S}}$, to search for the scene graph node, **n**, controlling the orientation of the head of the virtual character.
4. Attach an update callback with the node, **n**, to update the transformation matrix of the node to change the direction toward which the head looks to –*v*, every time the scene is rendered.
5. **end do**

This technique ensures the character always looks at the user. The rotation of the character's neck is constrained based on bio-mechanical joint angle limits of the human neck.

Note, that the above mentioned scenario is a typical AR scenario. The gaze control method will also work in a VR scenario where the viewpoint of the rendering camera can be controlled by mouse input from the user in real-time.

This method to compute the gaze of the character is entirely based on geometrical considerations. This makes the method simple to implement and very robust. It also makes the method efficient and suitable for real-time computation. We find that even this simple technique to compute gaze improves extent of immersion of the user in the virtual environment by a large extent.

4. Conclusion

In this work we provide a simple and efficient method for a real-to-virtual human camera-based gaze control, suitable for AR as well as VR applications. Based on the camera matrix provided by camera tracking (for AR) or user mouse input (for VR), we derive the inverse view direction of the virtual character's gaze, in order to establish a robust look-at behavior between virtual and real humans (as opposed to the real to virtual look earlier described in [4] and [8]).

However, the current method is still under development since more intuitive real-to-virtual camera controls need to be implemented to avoid certain problems, for e.g., the ego-motion of the real camera to be transferred to the virtual human or the real human looking in extreme angles with fast rotations of the head. Thus, camera stabilization methods to be used in conjunction with the camera tracking and geometrical damping for the virtual camera are parts of future work towards improving this method.

In this work, we aim to enhance 'believability' a term defined in [12] which is directly related to 'enaction.' An experience of realistic interaction in an interactive virtual environment (measured by believability) can instill an enhanced higher level of presence (which measures the feeling of 'being there') in the user. By providing an efficient and robust framework for animating virtual humans in virtual environments, where they have idle movements and body and face simulation during their interaction with real users [4] and by providing a simple gaze model in this work, we believe that both believability and presence will be significantly enhanced.

Acknowledgement

The work presented has been supported by the Swiss Federal Office for Education and Science in frame of the ENACTIVE NoE and the EU IST Programme, in frame of the INDIGO project.

References

[1] Argyle, M. and M. Cook. *"Gaze and Mutual Gaze"*, Cambridge University Press, New York, NY, 1976.

[2] Breton, G., Pele, D., Garcia, C., "Modeling Gaze Behaviour for a 3D ECA in a Dialogue Situation", IUI'06, ACM Press, 2006

[3] Ciger, J., Herbelin, B., Thalmann, D., "Evaluation of Gaze Tracking Technology for Social Interaction in Virtual Environments", *2nd Workshop on Modelling and Motion Capture Techniques for Virtual Environments, CAPTECH04,* Zermatt Dec. 9-11, 2004.

[4] Papagiannakis, G.,, Egges, A., Magnenat-Thalmann, N., *"Presence and Interaction in Mixed Realities"*, 3rd International Conference on Enactive Interfaces, November 20-21, Montpellier, 2006.

[5] Fitzgibbon A. and Zisserman, A. *"Automatic camera tracking"*, Video Registration (editors Shah, M. and Kumar, R.), pp.18-35, Kluwer Academic, 2003.

[6] Garau, M., Slater, M., Bee, S., Sasse, M., „*"The impact of eye gaze on communication using humanoid avatars"*, Conference on Human Factors in Computing Systems, SIGCHI01, ACM Press, 2001

[7] Kendon, A. *"Conducting interaction: patterns of behavior in focused encounters"*, Cambridge University Press, 1990.

[8] Papagiannakis, G., Magnenat-Thalmann, N., *"Mobile Augmented Heritage: Enabling Human Life in ancient Pompeii"*, International Journal of Architectural Computing, Multi-Science Publishing, issue 02, volume 05, pp.395-415, July 2007.

[9] Rothkopf, C., Pelz, J., "Head movement estimation for wearable eye tracker", Eye tracking research & applications ETRA'04, ACM Press, 2004

[10] Simon, G., Fitzgibbon, A. and Zisserman, A. *"Markerless tracking using planar structures in the scene"*, Proceedings of International Symposium on Augmented Reality, pp. 120-128, Oct. 2000.

[11] Sprague, N., Ballard, D., and Robinson, A. 2007. *"Modeling embodied visual behaviors"*, ACM Trans. Appl. Percpt. 4, 2, Article 11, July 2007

[12] Enactive Deliverable DRD2.3.1, "Guideline to the evaluation of ENACTIVE interfaces in terms of believability and preliminary recommendation to designers", 15 July 2006

Functional Analysis of Haptic Devices

Damien Couroussé Jean-Loup Florens
ACROE, Grenoble, France
E-mail: damien.courousse@imag.fr, jean-loup.florens@imag.fr

Abstract

This paper presents an attempt of classification of the different functional approaches one can find in the field of haptic research. Our methodology is based on a comparison of the natural interaction scheme involving a human and an object/environment with the very similar interaction scheme that exists in the mediated situation involving the haptic device. Four functional approaches emerge from our classification: Object, Human, Interaction and Task. The task-based approach is then divided into four sub-categories on the basis of topological criteria: Environment, Tool-Handled, Command and Object.

1. Introduction

Haptic science is situated at the intersection of several major disciplines of contemporary science: robotics, electronics, mechatronics, computer music, computer graphics, etc. Historically, haptic devices appeared relatively at the same time in the fields of robotics, teleoperation, computer music or HCI.

The large versatility of human gesture [9] is another important parameter. Indeed, human gesture involves many very different movements such as locomotion or dexterous manipulation, and manual tasks are used in very different aspects of human life. The versatility of human gesture and the broad range of human manual tasks cannot be embraced by only one haptic device, and there are currently as many forms of haptic devices as existing applications of haptics.

Haptics science has entered into a boiling phase, becoming a very complex and diversified field. Therefore marks for paving this particular field seem necessary to get an overview on what was done and to envisage the future. Considering the various domains from which haptic science comes from, the versatility of human gesture and the broad range of applications of the use of haptic devices, it appears that proposing a categorization of haptic devices is a difficult task.

Attempts of classification of haptic devices are found in large reviews of haptic technology and application fields [1], [4]. In these works, haptic devices are classified on the basis of two orthogonal dimensions: technology (grounded devices, exoskeletons, types of actuating and sensing, etc.) and application fields (medicine, telerobotics, etc.).

However, functional aspects of the haptic device are seldom considered. By functional, we understand the role or the function that the haptic device plays in the human-object interaction. We have found only one related work: in the paper presenting the FEELEX device [7], Iwata et al. propose a categorization of haptic devices on the basis of a functional approach. Three functional categories are presented: exoskeleton type force displays, tool-handling-type force displays and object-oriented-type force displays. In addition, several other approaches are laid apart: tactile display, passive input devices based on force sensing.

Our work aims at providing a functional overview of haptic devices in the many works and uses where we may find it. The second part of our paper details the methodology chosen. The third part constitutes a presentation of the different functional approaches that were extracted from our analysis.

2. Methodology

Our methodology is based on the comparison of the natural interaction with the mediated situation. It can be considered indeed that the haptic device plays the role of a medium between the human and the simulation "world" inside the computer [2]: this is what we call the "mediated situation".

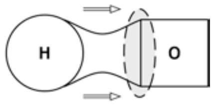
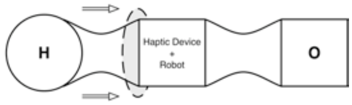
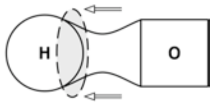
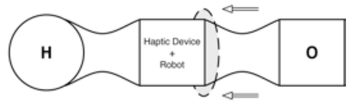
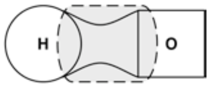
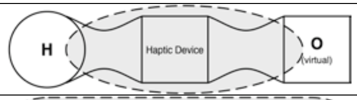
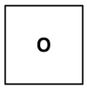
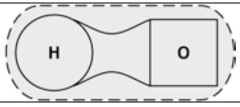
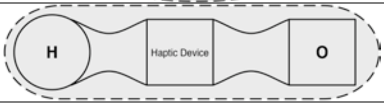
In the natural situation, the human interacts directly with an object or an environment thanks to mechanical coupling. In an artificial or mediated situation the aim of the haptic device is to reintroduce the mechanical coupling in the interaction. We can distinguish two different cases in the mediated situation. (1) A virtual object is "mediated" by the haptic device. (2) In the case of teleoperation, the object considered is real but manipulated through a haptic device. This situation is similar to the first one if considering the point of coupling between the human and the object.

3. Classification of functional approaches

3.1. Approach #1. Object

In this approach, the observer is interested in the properties of the object independently from those of the human. The hypothesis is made that the mechanical coupling can be cut so that the object can be studied alone. The object is considered as a separated entity

Table 1. Schematic representation of the different functional approaches in the analysis of human-object interaction

	Natural situation	Mediated situation	Legend	
#1 Object				Human
#2 Human				Mechanical Coupling
#3 Interaction				Object
#4 Task				Zone of focus & Point of view

from the human and is observed from its border limit with the human (Table 1 #1). In the mediated situation, the haptic device is often considered as a part of the object.

In this functional category we can find most of the reference analyses that attempt to delimit the conditions of acceptable behaviour for the haptic simulator. Transparency [8] and passivity [3] are two very important theoretical positions for haptics that apply to this functional category. Briefly speaking, one can say that they both try to define optimal conditions for the simulation of haptic objects.

In this functional category one can also cite some of the works for task simulation that are not oriented on the task but rather focus on the properties of the simulated object (e.g. simulations for surgery practice).

3.2. Approach #2. Human

In this approach, the observer is interested in the human and measures the physical human behaviour when involved in a task in interaction with a given environment (Table 1 #2). It is the symmetrical of the object-based approach. In the mediated situation the function of the haptic device is to transmit correctly the human action and behaviour to the object.

Similarly to the object-based approach, the point of observation is the border limit between the human and the object. The object-based approach is based on the measure of invariant properties such as stiffness, viscosity, mass or impedance. Conversely, the human is an active system with numerous DoF. Hence, the characterization problem becomes more complex, and the study of human action deals with motions and forces that can hardly be reduced to invariant properties.

The human-based approach is related to the field of psychophysics in general: the object of study of phsychophysics is the Human, and the mediated situation is considered as a mean to provide new stimuli. Two complementary positions exist:

1. The haptic device is be used to provide stimuli that would not or hardly be possible with common real objects [10].
2. The human is considered in the perspective of the mediated situation. Considered as a new situation for human perception and action, it becomes a case of study in itself [6].

3.3. Approach #3. Interaction

This approach focuses on interaction, by considering the human-object as a whole undividable dynamical system at its low mechanical level. The criteria induced on the function of the haptic device are in this case related to the comparison of two physical systems: (1) the natural one constituted of the human and the natural object and (2) the hybrid one constituted of the human and the mediated object (Table 1 #3).

In the natural situation, the “finger on the glass” is an exemplary case of this approach: the particular way the glass can “sing” comes from the very specific interaction (the slip-stick effect) that arises between the glass and the wet finger sliding on its border.

In the mediated situation, the haptic device is situated at the core of the interaction system that is observed. Indeed, the properties of the interaction system (and therefore the properties of the simulated object) are strongly related to the mechanical properties of the haptic device. This approach is related to the simulation of artificial situations where the quality of the human-object interaction plays a fundamental role. The general hypothesis is that the instrumental gesture, especially in the case of excitation gestures such as in violin bowing, must not be reduced to a simple control model from the player to the instrument [5].

3.4. Approach #4. Task

3.4.1. Introduction

In this approach, the observer is interested in the whole human-object system at many possible levels of observation (Table 1 #4). This level of analysis is the most general and includes all the very different interaction schemes that are not taken into account by the three previous categories of functional approaches.

The global methodology still consists in comparing the natural situation to the mediated one, but we introduce another level of analysis related to the topological relationships that exist between the human body, the object and other possible media like tools or environment.

3.4.2. Line of mobility

We assume that in every task there is necessarily at least one borderline of mobility along which the mechanical links are non-permanent during the duration of the task. This is what we call the line of mobility: it concerns the task, the mechanical objects and the linkages involved. The state of the line of mobility presents a temporal dimension as topological relations can evolve along time, but we can consider that there is a non-reducible amount of time during which the topological relationships will not evolve.

For example, when operating with a screwdriver during a session task that consists in unscrewing a screw the line of mobility is situated between the tool and the screw. When playing on a keyboard the line of mobility is situated between the hand and the keyboard.

3.4.3. Limit of the Virtual Environment

In the mediated situation, we need to examine the relationship of the line of mobility with another border: the limit of the virtual environment. Hence, a mediated object can have different status according to the point of view chosen [2]: (1) it can partially or totally become a part of the human body in the embodiment situation; (2) it can be a part of the environment. One can consider the two following extreme cases for reference of the limit of the VE:

"Restricted" virtual environment. The virtual environment is completely enclosed in the computation supported by the real-time simulator. In this case the haptic device is an interface that provides to the human operator the interaction access to this environment. Non-immersive environments often apply to this restricted definition.

"Enlarged" virtual environment. The virtual environment encompasses the physical separation between the human and the haptic device. This approach is used in spatial and immersive applications in which visual co-location is an important characteristic. The virtual environment may contain the operator's hands and then overlaps the real world.

3.4.4. Situation #a. Relation with a general environment

This situation concerns all tasks where the human link with the environment is not specified in the relationship with a particular type of object. In this case the line of mobility is necessarily situated at the limit of the human body (H) (Table 2-#a-left).

All kinds of navigation tasks are included in this category, such as body locomotion, but also more abstract kinds of navigations that include human action, such as sorting books in shelves.

The artificial situation is mediated by means of an exoskeleton that consists in a permanent link with the operators' body (Table 2-#a-right). It places the human in an immersive situation for at least a part of the degrees-of-freedom (DoFs) of the body, the degree of immersion depending of the part of the body attached to the exoskeleton. For example, arm exoskeletons are completed by hand or finger exoskeleton in order to support hand manipulation tasks.

3.4.5. Situation #b. Handled tool

In this situation a main line of mobility is located between the tool (T) and the operated object (E) since the link between the human's hand and the tool is supposed to be permanent all along the session task (Table 2.-#b). In the usual tool-object interaction two types of motions have to be considered that constitute the tool–object mobility:

- A "selection gesture", which consists in the evolution from non-contact to in-contact state (between T and E).
- An "ergotic gesture" [2], by which the tool (T) operates on the object or environment (E). Various relative motions and forces during the contact state are possible depending on the respective material and tribologic properties of the tool–object system.

The handled part of the tool is in the real world, whereas the non-handled part is in the virtual world, where it interacts with the virtual objects. Since the line of mobility is not situated at the border of the human body it can be easily included in the *restricted* virtual environment while a high level of transparency is not mandatory. On the other side since the dynamic system created by tool-object interaction is completely supported by the simulation software (the restricted VE), this type of configuration has to synthesize actions of complex tools including selection and contact motions.

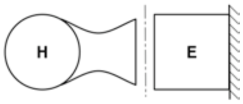
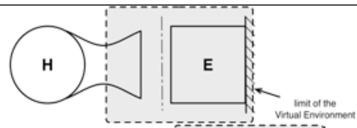
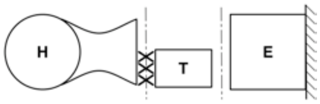
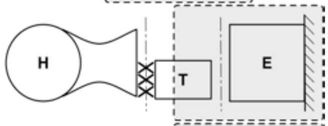
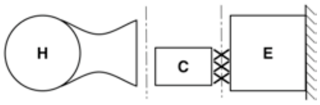
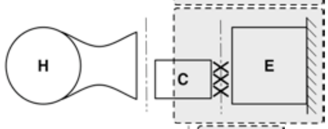
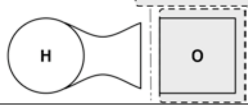
3.4.6. Situation #c. Actuating an artificial device

This category contains the hand actuating various types of artificial devices. This is usually done by the means of levers, buttons, handles, cranks, which we can categorize under the general term "command". In this situation the line of mobility is situated between the human (H) and the command (C) (Table 2-#c).

An example of this type of task is the manipulation of the gearbox lever. In this case the line of mobility is situated between the hand and the lever. Indeed the other parts of the mechanical chain from the lever to the ground are made of permanent joints and links. An other interesting case of this category is the piano keyboard playing.

In this situation the virtual environment does not necessarily contain the line of mobility. Indeed, the main difference with the situation #b is that the only

Table 2. The four categories of situations in the task-based functional approach

	Natural situation	Mediated situation
#a Environment / Immersion		
#b Tool-handled		
#c Command		
#d Object manipulation		

pertinent gesture around the line of mobility is the selection gesture and not the ergotic one. The ergotic gesture is performed in permanent grasping mode and the eventual sliding or local compression motions at handle-level have no incidence on the task.

3.4.7. Situation #d. Manipulating objects

This situation consists in direct interactions with objects. The line of mobility is situated between the human body and the object. The main difference with the situation #a is that objects are determined at the level of their interaction border by shape and possible limitations of the number of degrees of deformation. For example in the moving or actuating material objects, the interaction session may include contact and release phases. Like in the case of situation #b two main types of gestures can be distinguished: selection gestures that consist in free movements to reach a target object, and ergotic gestures associated to the contact phases.

In the mediated situation, the limit of the virtual environment includes the haptic device, and is concomitant with the limit of the object (Table 2.-#d). The gesture interface of the haptic device must present very particular properties because the line of mobility is *between* the human and the object.

4. Conclusion

We have extracted four main functional approaches of haptic devices through the analysis of theoretical positions, the design of new devices and applicative works. We have separated the task-based functional approach into four categories. These categories are not hermetically separated one to each other, but are at least differentiated on the basis of the topology of interaction, and on the limit of the virtual environment in the mediated situation.

This attempt of a functional analysis of haptic devices is a contribution to the numerous works reviewing haptic technology. We conceive it also as an opening into investigations and new analysis that could lead to a new approach or understanding of haptics.

References

[1] Burdea, G. C. (1996). Force and touch feedback for virtual reality. John Wiley & Sons, Inc.

[2] Cadoz, C. (2004). Enactive interfaces? Enactive Virtual Workshop. http://www.interdisciplines.org/enaction.

[3] Colgate, J. E. and Brown, J. M. (1994). Factors affecting the z-width of a haptic display. In Proceedings of the IEEE 1994 International Conference on Robotics & Automation, pages 3205–3210, San Diego, CA.

[4] ENACTIVE (2004). Report on the state of the art on existing sensors/actuators technologies for haptic interfaces. Deliverable D3.1, European Network of Excellence Enactive Interfaces.

[5] Florens, J.-L. (2003). Expressive bowing on a virtual string instrument. In Camurri, A. and Volpe, G., editors, Gesture-Based Communication in Human-Computer Interaction; 5th International Gesture Workshop, pages 487–496, Genova, Italy. Springer-Verlag Berlin Heidelberg 2004.

[6] Jansson, G. and Larsson, K. (2002). Identification of haptic virtual objects with different degrees of complexity. In Proceedings of Eurohaptics 2002.

[7] Iwata, H., Yano, H., Nakaizumi, F., and Kawamura, R. (2001). Project FEELEX: adding haptic surface to graphics. In Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pages 469–476. ACM Press.

[8] Lawrence, D. A. (1993). Stability and transparency in bilateral teleoperation. IEEE Transactions on Robotics and Automation, 9(5):624–637.

[9] Luciani, A., Evrard, M., Castagné, N., Couroussé, D., Florens, J.-L., and Cadoz, C. (2006). A basic gesture and motion format for virtual reality multisensory applications. In Proceedings of the 1st international Conference on Computer Graphics Theory and Applications (GRAPP), Setubal, Portugal.

[10] Tan, H. Z., Durlach, N. I., Beauregard, G. L., and Srinivasan, M. A. (1995). Manual discrimination of compliance using active pinch grasp: The rols of force and work cues. Perception & Psychophysics, 57(4):495–510.

Achieving Optimal Haptic Feedback in Telesurgery: Focusing on Transparency Alone is Not Enough

Koen Crommentuijn Dik J. Hermes Wijnand A. IJsselsteijn

Dept. of Technology Management, Subdept. of Human-Technology Interaction
Eindhoven University of Technology, The Netherlands
E-mail: k.j.crommentuijn@tue.nl, d.j.hermes@tue.nl, w.a.ijsselsteijn@tue.nl

Abstract

Surgeons are unable to use their sense of touch in telesurgery, because current telesurgical systems provide no haptic feedback. To overcome this problem, a lot of research is done on creating a surgical teleoperation system that does provide haptic feedback. Most of this research, however, focuses on the technology that needs to be developed. In this paper we advocate a different approach, considering at the same time; user, task and technology. With this approach, seven points of attention have been identified, which, if correctly addressed, should lead to an optimal solution of integrating haptic feedback into a telesurgical system.

1. Introduction

From 1990 on, laparoscopic surgery, also known as minimally invasive surgery (MIS), became a widely accepted and performed method of surgery [6]. In laparoscopic surgery, long slender instruments are used, generally with a diameter of 5 to 15 millimeters, to operate through small openings in the patient's body (see figure 1).

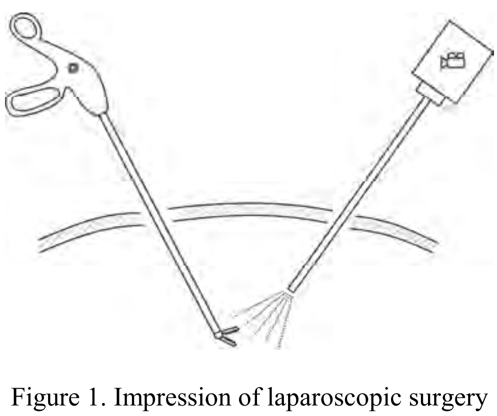


Figure 1. Impression of laparoscopic surgery

Laparoscopic surgery has many advantages compared with open surgery: Less blood loss, less post-operative pain, earlier return of bodily functions, shorter hospital stays, etc. [5]. Performing a laparoscopic procedure, however, is very demanding for the surgeon, and requires lots of skill and training. Main reasons for why laparoscopic surgery is more difficult than open surgery, are the loss of the natural view on the operative field, the limited freedom of movement of the laparoscopic tools, the reversed motion at the instrument tip due to a hinge point, the magnification of tremor through the long laparoscopic instruments, and the loss of direct touch [1, 5].

To solve the problems previously mentioned, new innovative technologies have been developed and introduced into the field of minimally invasive surgery. One such technological innovation was a robotic arm that can hold the laparoscopic camera and provide a steady view on the operative field. Later, dedicated surgical systems were built that could hold all the instruments. With these systems, the surgeon sits comfortably behind a console from where he controls the robotic arms holding specially designed instruments (see figure 2).

Figure 2. "da Vinci" Surgical System by Intuitive Surgical, Inc. (©2007 Intuitive Surgical, Inc.)

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

The application of teleoperation to the field of minimally invasive surgery is referred to as telesurgery. Currently, only the "da Vinci" surgical system, introduced in 1999 by Intuitive Surgical Inc., is commercially available and has been approved for a variety of surgical procedures by the American Food and Drug Administration (FDA). The "da Vinci" surgical system solves many of the problems related to laparoscopic surgery. A laparoscopic 3D-camera combined with a stereoscopic display give the surgeon a good 3-dimensional view on the operative field. New instruments have been designed that allow similar movements as the human wrist, increasing the freedom of movement of the instruments. The movements of the instrument tip are mapped into the same directions as the movements of the controls, solving the reversed motion issue. And, the natural tremor of the surgeon's hands is filtered out, resulting in a more steady and smooth movement of the instruments. Additionally, the movement of the instruments can be scaled down in relation to the controls, allowing for more precise and minute operations.

A surgical system as the "da Vinci" also has its disadvantages. The first disadvantage is of a financial nature; purchasing and maintaining a "da Vinci" surgical system is very expensive, especially when comparing the system with standard laparoscopic tools. The second disadvantage concerns the control of the surgical system. With the "da Vinci" surgical system, all haptic feedback, besides positional information, is lost. Without haptic feedback, surgeons have to rely completely on visual feedback to get an idea of the force they are applying via the robot. Surgeons are also unable to use their sense of touch to perceive important physical characteristics of the environment they are interacting with, which is important for identifying unhealthy tissue and finding reference points within the body.

Although surgeons express that they desire haptic feedback in a telesurgical system, the benefit of haptic feedback has not been empirically established (e.g. see [7]). A possible explanation is that the evaluated systems were limited in their capability to measure forces and used non-dedicated haptic interfaces to display the haptic feedback. Adding force sensors to surgical instruments and displaying those forces to the surgeon via generic haptic interfaces does not guarantee that the goals of providing haptic feedback will be achieved. Instead of focusing on haptic technology, we advocate a more human-centered approach; combining, (1) the requirements of the tasks and special environment belonging to minimally invasive surgery with, (2) the needs and abilities of the surgeon with, (3) the technology that needs to be developed (see figure 3). Currently, the question of how surgeons perceive haptic feedback via telesurgery, is virtually unexplored.

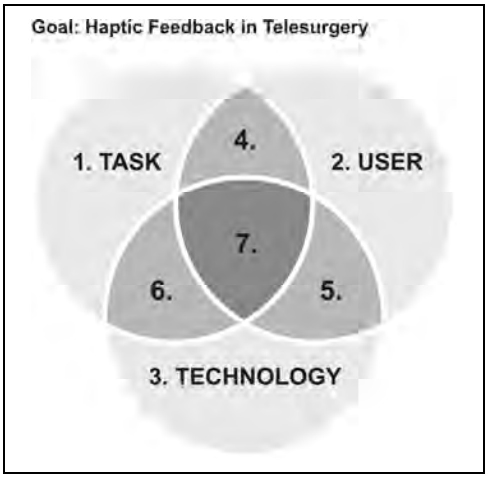


Figure 3. Combining task, user and technology

2. Combining task, user and technology

2.1. Task: Minimally invasive surgery

Minimally invasive surgery is a highly specialized domain with stringent requirements. Surgeons undergo years of training to become proficient, and new procedures and technologies are frequently introduced. To aid surgeons in their practice, it is necessary to understand the surgical procedures they perform and the special requirements that come with the field of minimally invasive surgery.

2.2. User: Surgeon

The surgeon needs to perceive relevant haptic information on the interaction between the surgical instruments and the environment. It is, therefore, necessary to understand how humans perceive their environment through physical contact. Concerning the sense of touch, a great deal on human physiology is understood. It is believed that all receptors contributing to the sense of touch have been identified and described. It is not well understood, however, how this sensory (afferent) information integrates with the higher-order processes, subsequently providing us with the perception of touch.

Physical variables, like friction, stiffness, viscosity and inertia are often described by simple equations. Stiffness, for instance, can be defined as the increase in resistive force over the deflection of a material. The sensitivity of humans to these physical variables has been examined in various studies (e.g. [3, 10]). It is still unknown, however, how these mathematically defined physical variables relate to the actual perception of friction, stiffness, viscosity and inertia. Even less is known about how physical properties lead to higher order percepts like roughness and texture. Bergman Tiest & Kappers [2], for instance, could not explain the dimensions they found in their extensive experiment in which participants had to sort different textures freely into groups. Because it is unknown how physical properties lead to higher order percepts, unexpected haptic effects have been found. Robles-de-la-Torre and Hayward [9], for instance, found that in perceiving a bump, the lateral force is more important than the geometric contour: When participants were presented with the geometric contour of a hole, accompanied by the lateral force of a bump, they generally perceived a bump.

2.3. Technology: Haptic feedback

The field of haptic feedback in teleoperation is still developing. New teleoperators with different characteristics are being and different ways of measuring forces are being investigated. One can, for instance, measure forces with dedicated sensors (e.g. by force/torque sensors or strain gauges), or measure forces indirectly (e.g. from the motor-input signal). Besides research done on hardware design, a vast amount of research is done on creating accurate and stable control models. A control model receives input from the control side and the effect side of the surgical system, and calculates the appropriate forces that need to be send back to both sides (see figure 4). There seems to be a trade-off between the transparency (i.e. how accurately forces at the instrument tips are displayed via a haptic interface) and the stability of the system. Many teleoperation systems that provide haptic feedback suffer from unwanted effects, for instance, when displaying a hard contact, or when a user grabs the controls firmly, a system can start oscillating. Needless to say, effects like these are unacceptable in a telesurgical system.

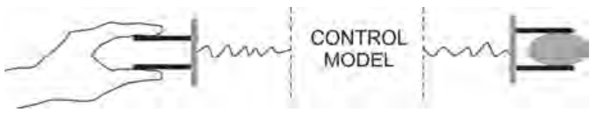


Figure 4. Impression of a teleoperation system

2.4. Combining task and user

The sense of touch provides the surgeon with feedback of his actions, for example, the amount of force he applies when dissecting tissue. Additionally, the sense of touch provides the surgeon with information on the physical properties of what he touches. In order to adequately provide haptic feedback to a surgeon, it is necessary to know how a surgeon uses his sense of touch during surgery. Questions like "how does a surgeon perceive that a knot is tied securely?", and "how does a surgeon identify different types of tissue, for example, a tumor?" need to be answered. We could start out by identifying and studying those tasks, in which surgeons desire haptic feedback most. In previous research, suturing and knot tying [4], dissecting [11] and detecting sclerosis [7] have been chosen as tasks of interest. What a surgeon needs to feel to successfully complete each of these tasks, however, has not been investigated.

2.5 Combining user and technology

Physical information is measured at the side of the surgical instruments and fed back to the surgeon. Haptic feedback thus depends on the type of information measured and the quality of this measurement. Haptic feedback is further degraded by the inevitable time-delay caused by measuring, processing and transporting information. Haptic feedback is additionally restricted by the haptic interface, which is limited in the reproduction of haptic information due to, for instance, not having force feedback in all degrees of freedom, internal friction, or the lack of a tactile display. To be able to make informed choices on the design and control of a telesurgical system with haptic feedback, it is necessary to investigate how technical limitations affect the perception of the surgeon (see for instance [8]).

2.6. Combining task and technology

A telesurgical system is a teleoperation device specially designed to meet the requirements of minimally invasive surgery. None of these requirements are to be violated when haptic feedback is incorporated into a telesurgical system. Instruments equipped with force sensors still need to fit through an opening of 5-15 mm. Additionally, sensors that could come into contact with the body of the patient need to be biocompatible, and to withstand the extreme conditions of a sterilization process.

It is assumed that a surgeon will make less force related errors when a telesurgical system is able to present haptic feedback. It is also possible, however, that more errors will be made because of haptic feedback. Care should be taken that a telesurgical system with haptic feedback remains 100 percent stable and no uninstructed or unintended movements are possible. Another problem related to force feedback is overshoot. For instance, a large force is sometimes necessary to drive a needle through tough tissue, but when popping out at the other side, the sudden drop in resistive force can cause an abrupt movement. Effects like these need to be investigated, to ensure that telesurgical systems with haptic feedback are, at least, just as safe as current telesurgical systems without haptic feedback.

2.7 Putting it all together

By combining task, user and technology, seven points of attention have been identified:

1. Knowing what the surgeon needs/wants to perceive about the physical interaction between the instruments and the environment.
2. Knowing how these percepts relate to measurable physical properties.
3. Designing the technology that is able to measure these physical properties, and complies with the requirements of minimally invasive surgery.
4. Designing the technology that is able to extract, and possibly enhance, the right information from measurements.
5. Designing a control model that is able to have the robot accurately follow the movements of the surgeon, and is able to accurately display the haptic information, while maintaining stability.
6. Designing a haptic interface that is able to display the haptic information in such a way that the surgeon perceives those things defined under point one.
7. Knowing how the limitations of the technology affect the perception of the surgeon, and taking this into account in the design of the system.

In order to design a telesurgical system with efficient haptic feedback, we believe that all seven points need to be addressed simultaneously.

Conclusion

Adding haptic feedback to a telesurgical system requires a human-centered approach. Current research on haptic feedback, however, generally focuses on transparency (i.e. how can forces accurately be copied from one side to the other, and vice versa). We believe that pursuing transparency alone, will not lead an optimal implementation of haptic feedback. Previous pursuits of transparency have, for example, resulted in systems that suffer from oscillations. By combining task, user and technology, seven points of attention have been identified. We believe that by addressing all seven points simultaneously an optimal integration of haptic feedback into a telesurgical system will be reached.

Acknowledgements

This research is funded by the IOP-MMI program of SenterNovem. The authors would like the thank Antal Haans for his contribution to this paper.

References

[1] Bann, S., Khan, M., Hernandez, J., Munz, Y., Moorthy, K., Datta, V., Rockall, T. "Robotics in Surgery", Journal of the American College of Surgeons, vol. 196(5), pp. 784-795, 2003.

[2] Bergmann Tiest, W.M., Kappers, A.M.L. "Analysis of haptic perception of materials by multidimensional scaling and physical measurements of roughness and compressibility", Acta Psychologica, vol. 121, pp. 1-20, 2006.

[3] Jones, L.A., Hunter, I.W. "A perceptual analysis of stiffness", Exp. Brain Res., vol. 79, pp. 150-156, 1990.

[4] Kitagawa, M., Okamura, A.M., Bethea, B.T., Gott, V.L., Baumgartner, W.A. "Analysis of Suture Manipulation Forces for Teleoperation with Force Feedback", Proc. of the Fifth International Conference on Medical Image Computing and Computer Assisted Intervention, vol. 2488, pp. 155-162, 2002.

[5] Lanfranco A.R., Castellanos, A.E., Desai, J.P., Meyers, W.C. "Robotic Surgery: A Current Perspective", Annals of Surgery, vol. 239(1), pp. 14-21, 2004.

[6] Litynski, G.S. "Endoscopic Surgery: The History, the Pioneers", World Journal of Surgery, vol. 23(8), pp. 745-753, 1999.

[7] Mayer, H., Nagy, I., Knoll, A., Braun, E.U., Bauernschmitt, R., Lange, R. "Haptic Feedback in a Telepresence System for Endoscopic Heart Surgery", Presence, vol. 16(5), pp. 459-470, 2007.

[8] O'Malley, M., Goldfarb, M. "The Effect of Force Saturation on the Haptic Perception of Detail", IEEE/ASME Transactions on Mechatronics, vol. 7(3), pp. 280-288, 2002.

[9] Robles-De-La-Torre, G., Hayward, V. "Force can overcome object geometry in the perception of shape through active touch", Nature, vol. 412, pp. 445-448, 2001.

[10] Tan, H.Z., Durlach, N.I., Beauregard, G.L., Srinivasan, M.A. "Manual discrimination of compliance using active pinch grasp: The roles of force and work cues", Perception & Psychophysics, vol. 57(4), pp. 495-510, 1995.

[11] Wagner, C.R., Stylopoulos, N., Howe, R.D. "The Role Of Force Feedback In Surgery: Analysis Of Blunt Dissection", Proc. 10th Symp. on Haptic Interfaces for Virtual Env. & Teleoperator Syst., pp. 68-74, 2002.

Exploitation of manual skill for the generation of aesthetic shapes: a typical enactive approach

Umberto Cugini and Monica Bordegoni
Dipartimento di Meccanica, Politecnico di Milano, Italy
E-mail: umberto.cugini@polimi.it; monica.bordegoni@polimi.it

Abstract

This paper presents the results of a research work aiming at developing a system for the generation of aesthetic shapes that exploit the manual skill of model makers and stylists. The analysis of model makers' activities in the real workshop has allowed us to identify the role and importance of the sense of touch in shape evaluation and to design the system based on haptic interfaces. The paper presents two haptic interfaces for virtual shape evaluation: a sand tool resembling sand paper and a haptic strip resembling "red tape" typically used in the automotive field.

1. Introduction

Gestures, besides speech, represent the mostly used means of expression by humans. For what regards the product design field, designers have multiple ways for communicating their ideas and concepts. One of them concerns *model making* activity, where designers make explicit their concepts by using some appropriate tools and specific hand movements on various material with the intent of obtaining a shape. In this process they develop specific high skills and a personal "language".

Some studies have demonstrated that visual, tactile and kinesthetic feedbacks are jointly important in shape creation and evaluation process [14]. Specifically, a craftsperson needs to touch his work that is the physical mock-up. This touch is physical and continual, and provides control of the whole shape creation process [12], which is a creativity process. The process is a dynamic and evolving: stylists give form to shapes by continuously interacting with them, thus operating in a continuous loop of creation and perception activities.

Modelers gather plenty of information by touching the physical mock-up, encompassing it, sliding hands around it, manipulating it and, of course, simulating a real use. Modelers take benefit of these feedbacks to create shapes, to comprehend them, to check their adequacy to the initial or evolving idea and some times to the preliminary definition in form of drawings, and to evaluate the mock-up quality. By gestures, they enrich their mental representation of the shape, which help to appreciate and criticize its features.

Humans own knowledge on concrete actions in hands. Specifically, in the industrial design field, modelers own knowledge on manually sculpting malleable materials. *Knowledge* is physical, experiential, personal and contextual. This knowledge is also named know-how, operative, skill, *enactive*. Skill can be defined as the learned ability to do a useful process well, and is acquired by demonstrational and sharpened by practice. It is a personal and company's asset that requires to be preserved and exploited.

Model making practice does not allow stylists and model makers to directly create a digital model (virtual prototype) of the object being created and necessary for the downstream product development phases. Conversely, current design tools (CAD/CAS) allow the generation of digital models but do not offer functionalities that are oriented to designers' abilities, competences and skills.

The objective of the research work described in this paper, partially supported in the frame of two European funded projects (*T'nD* and *SATIN*), is building some tools for shape modeling that although based on direct modeling of a digital shape, allows stylists to work and interact with the digital model with their hands and through some *haptic tools* thus preserving and exploiting their skill.

2. Related works

The topic related to interactive shape modeling has been widely addressed in literature. Some works have proposed the use of sketches and strokes for creating shapes in the 3D dimensional space [16]. Other works are based on the inference of a 3D model from sketched curves [10]. These systems do not provide any physical contact with the created digital shape.

A survey of virtual clay for interactive modeling has been recently presented in [3]. General models

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

based on geometric and volumetric support global deformations, and mimic the global and local effects of plasticity, mass preservation and surface tension that are typical of clay material [7].

Some other applications are more related to physics-based shape modeling. An overview of physics-based deformable models is provided in [13]. Some sculpting systems have been developed based on haptic force associated with dynamic subdivision of solids which give users the illusion of manipulating semi-elastic virtual clay [4, 6].

Existing haptic devices are mainly used for clay modelling and are not specifically oriented to the industrial design sector. Most of them are based on the use of the point-based Phantom stylus by SenSable Technologies Inc. (http://www.sensable.com). The only physically-based shape modeling system commercially available is FreeForm, by SenSable Technologies Inc., which is based on the Phantom haptic device. Users work directly with the digital clay using the Phantom stylus as a modeling tool. Therefore, the interaction happens through a point contact, providing very limited information about the surface perception and control of the global geometric characteristics.

In conclusion, most of the works presented support local deformation of the model, and the haptic device is used to modify voxel or polygonal mesh models that are mainly oriented to applications for virtually sculpting small objects where the shape is controlled in a qualitative way. Conversely, the systems described in this paper address the industrial design domain satisfying precision and surface quality constraints, and allow users to explore object surfaces by means of intuitive haptic interaction tools and modalities [1].

3. Types of hand interaction and gestures

In order to achieve the goal of proposing an intuitive and easy-to-use shape modeling system, the research activity supported by cognitive psychologists participating to the research project has focused on the observation and analysis of model makers' activities performed in the real workshop [5].

The intention was to gather information about the different modeling techniques used by modelers and stylists at work, their gestures, the used tools, and the modalities used for checking the quality and characteristics of the in-progress models. Therefore, practices and hand modeling activities performed by modelers of the projects industrial partners (working in the car design and domestic appliances design sectors) have been observed and analyzed. A quantitative analysis of gestures (characterized by the kind of material used, the modelers' skills, the expected level of accuracy and the overall complexity of the mock-up) has been performed. Main gestures have been selected and described on the basis of qualitative and quantitative information, such as: aim and modus operandi of the tool, tool and hands movements, shape evaluation tasks, similarity with other tools, etc.

We can identify two main types of gestures: ergotic gesture which is performed by the modeler when working on the shape [2], and exploratory gesture which is performed for evaluating the quality of the shape [11].

Ergotic gestures and tools

The analysis of the users' modeling sessions has highlighted that 75% of the modeling actions are done using few tools. The movement performed for scraping material is representative enough of most of the gestures families occurring during the process. This action is rather qualitative, and the same gesture is used both for shaping and finishing a surface, and is also used either for large surfaces or for details. Concerning the type of tools, rakes are used for removing thick layers of material, while sandpaper is used for finishing and smoothing the surface and/or to obtain a uniform level of roughness.

Exploratory gestures

After working the piece of material with the tools, the modeler checks whether the physical mock-up has the expected proportions and quality of shape. Diverse approaches are used for assessing the correctness of the work. Modelers are used to check the object dimensions, the profile and the curvature continuity through the use of both vision and touch. These evaluations depend on the expected level of accuracy, on the scale and size of the mock-ups, and on personal strategy of the modeler. These actions are used to test the final shape (sculpting steps), and to ensure the obtaining of high quality surface (finishing steps).

Tactile verifications are carried out in different ways, each of them having different goals. Rapid sweeps on surface appears often while sculpting or finishing the mock-up. This action carries information about the detection of irregularities or variations on the curvature. This method allows the detection of features otherwise unrecognizable by means of visual observation only. Long sweeps are performed along the object surface. These are iterative movements, back and forth or from side to side, through which the modeler gets information about curvature variations more than absolute curvatures themselves.

In conclusion, tactile and kinesthetic inputs seem to complete visual information, still ambiguous about the real size and dimension of objects. This information can be only acquired through kinesthetic systems exploring the whole objects with the movement of our arms. This exploration helps constructing a more precise 3D mental representation of the conceptual shape.

4. Haptic perception

The sense of touch can discriminate and recognise complex objects although it tends to respond best to different aspects of objects than the visual system does [9]. When we move our limbs about actively through the world, we perceive objects through a combination of tactile and kinesthetic sensations caused by our mechanical interaction with them. Such experiences play an important role in perceptual development as vision and touch calibrate each other, and they can also be important under conditions in which visual and auditory information about the world is missing or impoverished. Our experience of the world based on a combination of tactile and kinesthetic sensation is called haptic perception [8].

The term *haptic perception* refers to perceptual processing of inputs from multiple subsystems, including those in skin, muscles, tendons, and joints. Haptic perception is usually active and information-seeking: the perceiver explores the world rather than passively receiving it. We can distinguish between passive and active touch. *Passive touch* provides tactile information; while *active touch* combines tactile and kinesthetic information can involve varied number of fingers and body parts and involve active exploration of the properties of objects (position, size, weights, texture, color, etc.) as well as object recognition or haptic search tasks.

People are very good at identifying ordinary objects presented haptically. Our ability to discriminate objects size and length using only haptic perception is quite good. Several other haptic attributes of objects can be reliably discriminated and are often used in haptic object identification: surface texture and perception of hardness and curvature of the object surface [15]. Lederman & Klatzky [11] coined the term *exploratory procedure* for a particular way of feeling an object in order to extract one or more of its properties.

The work here described aims at developing haptic tools to perceive surface curvatures and curvature variations. Studies and tests have been performed to understand human capabilities in perceiving surface properties. They have shown that the human capability in perceiving surface properties such as discontinuities and curvature is not so effective using haptic contact only, while that is well accomplished through the combination of vision and touch. This information is useful for defining the haptic interface specifications in terms of resolutions and performances.

5. Haptic tools for evaluating shapes

Objective of our research is to study and develop haptic tools for surface exploration with the aim of checking the surface quality. Two haptic tools have been studied: sanding tool and haptic strip.

Sanding tool

The FP6-IST-2002-001996 T'nD – Touch and Design project (http://www.kaemart.it/touch-and-design) has implemented an innovative virtual clay modeling system based on novel haptic interaction modality oriented to industrial designers. The project has developed a sanding tool supporting double curvature that allows the simulation of a piece of sandpaper, where the curvature actually follows the curvature of the virtual surface to be sanded. This allows for an accurate and realistic representation of the surface shape. The haptic tool simulates the underlying surface of the work piece contacted, which has curvature in addition to orientation. Our system is limited to the G2 curvature, which is the degree of "roundness" of the surface, similar to the second derivative of a curved line on a flat surface. From mathematical theory, we know that the curvature has three degrees of freedom:

- Orientation of the first principal direction.
- Curvature in the first principal direction.
- Curvature in the second principal direction, which is orthogonal to the first.

According to that, the hardware device consists of a plane structure based on semi-stiff radial ribs placed at 120° intervals that is forced by the actuators to the desired cross-sections of the surface to the desired curves (Figure 1a). The structure is covered by a tactile patch. The patch is supported and controlled from the centre, where the G0 location, G1 attitude orientation and G2 curvature of the patch are defined and controlled. A detailed description of the tool is reported in [5]. The sanding tool is easily manipulated by the user as it was a real piece of sandpaper (Figure 1b). The user is free to move and rotate the tool around the three axes in order to feel the surface curvature of the object. When the user moves the hand the device deforms according to the surface curvature.

a) b)

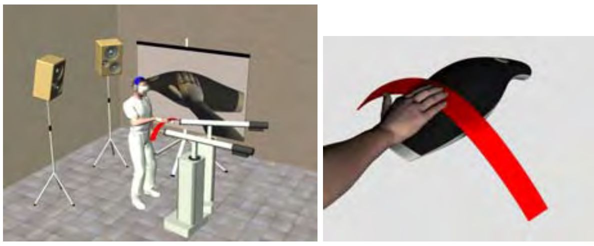
Figure 1. Sanding tool for shape quality evaluation. a) Tool structure; b) Tool in use.

The evaluation of the functional aspects and the performances of the tool developed reported that due to some hardware limitation the users were not able to feel small details of the model, but they were indeed able to feel big variations in curvature. Then, we have launched a new project for developing a new concept of haptic interface for shape evaluation.

Haptic strip

An innovative system for exploring and modifying shapes is being studied and implemented within the framework of the European project FP6-IST-5-034525 SATIN– Sound And Tangible Interfaces for Novel product design (http://www.satin-project.eu). The goal of the project is to develop and test innovative interfaces that will be used by designers for exploring and evaluating, and also modifying digital shapes through free-hand interaction that allows the exploitation of designers' existing skills. The project is based on new multimodal interaction paradigms where visual, haptic and sound modalities are fused together in order to provide designers with information about shape properties in response to evaluation and modification activities. The SATIN system consists of an Augmented Reality environment where the user can see 3D shapes, touch those shapes and also explore multiple geometric properties of shapes by means of metaphoric sounds (Figure 2 a).

The system will consist of a tangible interface allowing the evaluation, control, and modification of shapes. In order to "feel" the shape, the user conceptually leans a tape over a virtual surface. The tape is tangent to the surface along a line and physically creates a tangible path along the exploring trajectory. The user passing his hand on the tape feels and physically perceives the shape.



a) b)

Figure 2. a) Conception of the SATIN system. b) haptic strip for shape exploration.

Some tests on human haptic perception of shapes have provided information useful for the specification of the resolution of the haptic strip. In fact, due to technological limitations in the haptic interface development some surface features cannot be represented (for example, discontinuities). Therefore, some other modalities, such as metaphoric sounds, are used for conveying surface properties e.g. distance clearance, closeness to a curvature, or torsion limit that cannot be rendered haptically.

6. Conclusions

This paper has presented novel systems based on haptic interfaces for exploring and checking surface quality of industrial design products. The systems adopt typical enactive approach, in the sense that they exploit manual skill of stylists and model makers for the generation and evaluation of aesthetic shapes. The approach has been evaluated as effective from the final users of the systems, and fully in line with their typical way of working.

References

[1] Bordegoni M. and Cugini U., A conceptual design tool based on innovative haptic devices, Proc. DETC'06, ASME, Philadelphia, 2006.

[2] Cadoz C., Wanderley M.M., Gestures in Music, In M. Wanderley and M. Battier (eds.) Trends in "Gestural Control of Music". Paris, Fr: IRCAM - Centre Pompidou, 2000.

[3] Cani M. and Angelidis A., Towards virtual clay, in course "Interactive Shape Editing", SIGGRAPH, 2006.

[4] Celniker G. and Gossard D., Deformable curve and surface finite-elements for freeform shape design, Computer Graphics, 25, 4, 257-266, 1991.

[5] Cugini U. and Bordegoni M., Touch and design: novel haptic interfaces for the generation of high quality surfaces for industrial design, The Visual Computer Journal, Springer, 23(3), 2007.

[6] Dachille F., Qin H., and Kaufman A., A novel haptics-based interface and sculpting system for physics-based geometric design. Computer-Aided Design, (33), 2001.

[7] Dewaele G., Cani M.P., Virtual clay for direct manipulation, Eurographics, 2004.

[8] Gibson J.J., The senses considered as perceptual systems, Boston: Houghton Mifflin, 1966.

[9] Klatzky R.L., Lederman S.J. and Metzger V., Identifying objects by touch: An "expert system,", Percept Psychophys 37: 199-302, 1985.

[10] Igarashi T., Matsuoka S., and Tanaka H., Teddy: a sketching interface for 3D freeform design, in SIGGRAPH'99, ACM Press, 1999.

[11] Lederman S.J. and Klatzky R.L., Hand movements: a window into haptic object recognition, Cogn Psychol 19:342-368, 1987.

[12] McCullough M., Abstracting craft – The Practiced Digital Hand, MIT Press, 1998.

[13] Nealen A., Mueller M., Keiser R., Boxerman E. and Carlson M., Physically-based deformable models in computer graphics, in State Of the Art report, Eurographics, 2005.

[14] Poitou J.P, La dissonance cognitive, A.Colin, Paris 1974.

[15] Pont S.C., Kappers A.M.L. and Koenderink J.J., Haptic curvature discrimination at several regions of the hand, Perception & Psychophysics, 59, 1225-1240, 1997.

[16] Wesche G. and Seidel H., FreeDrawer: a free-form sketching system on the responsive workbench, in Proceedings of ACM Symposium on Virtual Environments, 2001.

Developmental approach of postural control from 7 to 11 years old and adults when proprioceptive inputs were disturbed

Rémy Cuisinier[*] Isabelle Olivier[*] Marianne Vaugoyeau[†] Vincent Nougier[*] Christine Assaiante[†]

(*)Université J. Fourier, Laboratoire TIMC-IMAG, CNRS, UMR 5525, Equipe « Santé, Plasticité, Motricité », F-38000 Grenoble, France

(†)UMR CNRS 61 96, Groupe "Développement et Pathologie de l'Action", Université de la Méditerranée, 31 chemin Joseph Aiguier, BP 71, 13402 Marseille cedex 20, France.

E-mail : Remy.cuisinier@ujf-grenoble.fr

Abstract

The study examined, in children aged 7 to 11 and adults, the effects of a proprioceptive perturbation in the control of posture when eyes were opened. To achieve this goal, 37 children aged from 7 to 11 years and adults were asked to stay as more stable as possible on a force platform in two conditions of no vibration and vibration of the ankle joint. The results 1) confirmed a non monotonic pattern of postural control development from 7 years old, and 2) showed that proprioceptive perturbation affected children from 7 to 11 years old but not adults suggesting that other developmental phenomenon in relation with the integration of proprioceptive inputs appeared after 11 years old.

1. Introduction

The integration of visual somatosensory, and vestibular inputs for the control of posture develops throughout childhood [2,7]. Improvement of postural control in children has been described as a decreasing magnitude of postural sway of the center of foot pressure (CoP) [3,4,5]. Maintaining bipodal equilibrium suggests that subjects behave like an inversed pendulum. By using such a biomechanical model, a close correlation was found between the displacement of the CoP and the subjects' center of mass. Thus the CoP provides information about the control process or strategy adopted to keep stance [13]. Under altered sensory conditions (eyes closed and/or ankle joint inputs altered), researchers found a shift in controlling sensory inputs to posture with development. This shift resulted from a visual dependence in the 4- to 6-year-old to a more adult-like dependence involving a combination of ankle joint and visual inputs [7]. These authors suggested that the adult form is reached between 7 and 10 years old, and improves again after 7 years old [8]. However, it is proposed that the age 7 is a transition period in the development of postural control [7,9], suggesting that the improvement of postural control throughout childhood is non-monotonic.

Even if vision is dominant for postural control, proprioceptive inputs play an important role. Several experimental approaches for investigating the role of proprioceptive inputs in postural control are available. One consists in applying vibration to muscles tendon [1,6,11]. The mechanism of this stimulation is based on a selective activation of muscles spindles, predominantly Ia afferents [12]. When vibration is applied, the central nervous system receives a disrupted muscle-spindle message, increasing subjects' postural sway. However, in order to test the role of proprioception in postural control, instructions are usually to stay as stable as possible and to keep eyes closed because visual cue is major to control posture.

This study investigated the development of postural control from 7 to 11 and adults when proprioceptive inputs were disturbed with opened eyes. We therefore conducted a study of quiet stance to determine whether non-monotonic events would pertain and whether disturbed proprioceptive inputs increase postural sway at all age in the same way even when vision is allowed.

2. Method

2.1. Subjects

37 children of 7 years old (N= 8, M=7.3 years ± 3 months), 8 years old (N= 8, M=8.1 years ± 1 month), 9 years old (N= 7, M=9.2 years ± 4 months), 10 years old (N= 6, M=10.1 years ± 2 months), 11 years old (N= 8, M=11.4 years ± 3 months), and 9 adults (M=25.7 years ± 27 months) participated in the experiment. All participants were right-handed and

naive as to the purpose of the experiment, had a normal scholastic level and did not show any known neurological or motor disorders. This study was approved by the local ethics committee and, in conformity with the Helsinki Convention, informed consent was obtained from all subjects.

2.2. Procedure

Subjects, arms close to the trunk, stood barefoot in a bipodal stance position on the force platform (AMTI®, model OR6-5-1). In order to increase body instability, their feet were placed slightly apart on marks drawn on the platform in the semi-tandem position with the right foot before the left one (cf. figure 1). Signals from the force platform were recorded along the vertical (F_z) direction and the torque was measured in the frontal (M_y) plane (100 Hz frequency with a 12 bit A/D resolution).

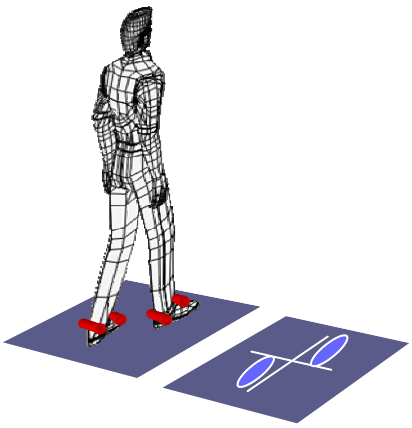


Figure 1. Experimental set-up: note the semi-tandem position with two vibrators (in red) disturbing the proprioceptive signals from the ankle.

The proprioceptive perturbation was delivered by two mechanical vibrators (280 g, diameter 4 cm, length 8 cm), which consisted of biaxial motors equipped with small eccentric masses rotating at 80 Hz and delivered a 3 mm amplitude. The vibrators were attached to the subjects' ankles with rubber bands on the tendons of both triceps surae muscles and on the tendons of both tibialis anterior.

Two blocks of four trials were executed (with and without proprioceptive perturbation). Whatever the block, a trial had a duration of 30 s. Subjects were instructed to look straight ahead and to stay as stable as possible.

Force plate data were filtered with a 10 Hz low-pass, second order Butterworth filter. Because subjects were placed in a semi tandem position, only displacements of the centre of foot pressure (CoP) in the medio-lateral plane were calculated using the following approximation: $\Delta CoP = \Delta My/Fz$ in which ΔMy was a change of the torque with respect to its baseline value (defined as the average value within the time interval from 0 to 30 s). Thus, we calculated the mean amplitude (MA) and velocity (MV) of CoP displacements in the medio-lateral direction.

2.3. Statistical Analysis

In order to explore the effects of proprioceptive perturbation through age, a 6 Ages (7, 8, 9, 10 ,11 years and adults) x 2 levels of proprioceptive perturbation (non-Vib and Vib) analysis of variance (ANOVA) with repeated measures on the last two factors was applied to MA and MV. The Tuckey HSD Post-hoc test was used whenever necessary. The level of significance was set at p<0.05.

3. Results

Analysis of MA showed a tendency to a main effect of age $F_{(5,40)}$=2.34, p=0.057. The Mean Amplitude of the CoP decreased from 7 years to adults (5.46 ±1.45 mm vs 4.06±1.08 mm, respectively). The MA did not vary between 8, 9, 10, 11 years and adults (4.51±1.30 mm 4.40±1.70 mm 4.62±1.02 mm 4.44±0.52 mm and 4.06±1.08 mm, respectively). The MA of CoP seemed to be more affected in the 7 years old children.

In addition, this analysis showed a main effect of proprioceptive perturbation $F_{(1,40)}$=.75, p<0.005. Indeed, when this sensory perturbation was applied, the MA increased from 4.59±1.12 mm to 5.14±1.18 mm.

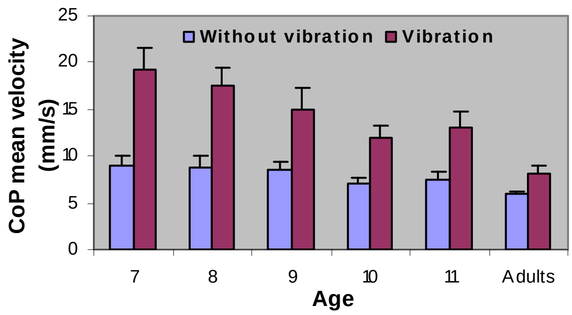


Figure 2. Evolution of the mean velocity of the CoP (mm/s) as a function of the proprioceptive perturbations with respect of age.

As illustrated in Fig. 2, analysis of MV showed a significant two-way interaction of Age x Proprioceptive perturbation, $F_{(5,40)}$=9.25, p<0.0001. Tukey HSD post-hoc test revealed that no difference reached the threshold of significance whatever the age when no proprioceptive perturbation was applied. As opposed to the proprioceptive perturbation condition, the mean velocity of the CoP significantly diminished from 7 years to 10 years and to adults, respectively. In addition, this velocity increased for all ages excepted in adults when the proprioceptive perturbation was applied.

4. Discussion

The primary findings of this study were the following:

First, from 8 to 11 and adults, the mean amplitude of the CoP did not vary. This result confirmed that there was a modification in the postural strategy from 7 to 8 years old as a function of proprioceptive perturbation. Whereas the evolution of MA throughout childhood was non-linear (i.e, MA diminished from 7 to 8 years old who reached the level of adults), MV decreased linearly from 7 to 10 years old. This result confirmed the presence of a non- linear pattern of postural control development in children aged from 7 to 8 years old [5,7,9]. 7 years old were more perturbed than 8 years old when proprioceptive information was disturbed.

Second, results based on the mean velocity of the CoP showed that disturbed proprioceptive inputs affect postural control only during childhood when eyes were opened. Indeed, the mean velocity of the CoP decreased from 7 to 11 years old. Children of 11 years old did not reach the adults' level. This result suggested that 1) even if eyes were opened, proprioceptive inputs were useful and integrated by the central nervous system to control posture during childhood, and 2) other developmental phenomenon appeared after 11 years old [10]. In adults, the mean CoP velocity did not seem to be perturbed by disturbed proprioceptive inputs when eyes were opened whereas the mean amplitude did. This particular sensitivity to vibration in adults suggested that they used a different strategy than children of 11 years old when visual inputs were available (i.e. a decreased velocity of postural sway with the same amplitude).

Further research is thus needed to extend the population to teenage in order to determine the age to which the postural strategy looks like adults form.

5. References

[1] Eklund, G. “General features of vibration-induced effects on balance”. Ups. J. Med. Sci. 77, pp. 112-124, 1972.

[2] Forssberg, H., Nashner, L.M. “Ontogenetic development of postural control in man: adaptation to altered support and visual conditions during stance”. J. Neurosci. 2, pp. 545-552, 1982.

[3] Hayes, K.C. “The development of feed forward control of podture in children”. PhD thesis, University of Waterloo, Waterloo ON, Canada, 1985.

[4] Hayes, K.C., Riach, C.L. “Preparatory postural adjustments and postural sway”. In: Woollacott MH, Shumway-Cook A (eds) Development of posture and gait across the lifespan. University of South Carolina Press, Columbia, South Carolina, pp. 95–127, 1989.

[5] Kirshenbaum, N., Riach, C.L., Starkes,J.L. “Non-linear development of postural control and strategy use in young children: a longitudinal study”. Exp. Brain Res. 140, pp. 420-431, 2001.

[6] Lackner, J.R., Levine, M.S. “Changes in apparent body orientation and sensory localization induced by vibration of postural muscles: vibratory myesthetic illusions”. Aviat. Space Environ. Med. 50, pp. 346-354, 1979.

[7] Shumway-Cook, A., Woollacott, M.H. “The growth of stability: postural control from a developmental perspective”. J. Mot. Behav. 17, pp. 131-147, 1985.

[8] Riach, C.L., Hayes, K.C. “Maturation of postural sway in young children”. Dev. Med. Child. Neurol. 29, pp. 650-658, 1987.

[9] Riach, C.L., Starkes, J.L. “Velocity of center of pressure excursions as an indicator of postural control systems in children”. Gait Posture 2, pp. 167-172, 1994.

[10] Sparto, P.J., Redfern, M.S., Jasko, J.G., Casselbrant, M.L., Mandel, E.M., Furman, J.M. “The influence of dynamic visual cues for postural control in children aged 7-12 years”. Exp. Brain Res. 168, pp. 505-516, 2006.

[11] Roll, J.P., Vedel, J.P. “Kinaesthetic role of muscle afferents in man, studied by tendon vibration study”. Exp. Brain Res. 47, pp. 177-190, 1982.

[12] Roll, J.P., Vedel, J.P., Ribot, E. “Alteration of proprioceptive messages induced by tendon vibration in man: a microneurographic study”. Exp. Brain Res. 76, pp. 213-222, 1989.

[13] Winter, D.A. “Biomechanics and motor control of human movement”. 2nd ed, John Wiley & Sons, New York, 1990.

Feeling objects in Virtual Environments: Presence and Pseudo - Haptics in a Bowling Game

Kristina Daniliauskaite
kdan05@imi.aau.dk

Ágústa Magnúsdóttir
amag05@imi.aau.dk

Henrik Bjørkå
hbbj05@imi.aau.dk

Lucas A. Hansen
laha05@imi.aau.dk

Jesper Pærregaard
jpae05@media.aau.dk

Luis E. Bruni
leb@imi.aau.dk

Stefania Serafin
sts@imi.aau.dk

Aalborg University Copenhagen
Medialogy

Abstract

Interrelations between the senses may be a central psychological component when it comes to producing the sense of presence in a virtual environment (VE).

In this paper we report the design and implementation of a VE in which the user registers the sensation of a haptic property - weight, by relying on visual cues, taking therefore advantage of sensory substitution (no haptic feedback device is actually present).

The interdependency between presence and a pseudo-haptic feedback is investigated by building a virtual bowling game. Results indicate that there is a significant correlation between the sensation of presence and the occurrence of pseudo haptics.

1. Introduction

The interfaces implemented in today's VE often engage only few of the senses, such as aural and visual. By stimulating more channels of the human sensorial system (i.e. haptic, proprioceptive, and olfactory) the user could gain a perceptually more detailed experience of the VE, which in turn could enhance the user's sense of being present.

An approach to stimulating more channels is to make use of knowledge from on-going cross-modal perception research. Certain perceptual illusions are an instance of cross modality that can be created through the alteration of a normal sensory perception, where one modality governs over another; creating the illusion. The illusory phenomenon has in some cases been proven to cross over modalities as a form of synesthesia where the modality in question is not being stimulated at all [2]. In this case sensory substitution can be used, employing sensory channels such as visual to stimulate the haptic sense.

1.1. Pseudo - haptic feedback

Empirical research verifies that visual feedback and a passive input device used together can evoke a pseudo - haptic feedback [4].

Recent investigations show that tactile properties of an object (i.e. texture and weight), can be obtained by the use of the control/display (C/D) ratio technique. This is the relation between a control device (e.g. a mouse) and a correlating object on the screen display (e.g. a cursor).

Lécuyer et al. [4] investigated how manipulation of the C/D ratio generates a haptic sensation of textures. Acceleration of a cursor, moving faster, implies a negative slope, such as going downhill. Similarly, deceleration, a positive slope in the texture, would indicate an uphill climb. Thus, the speed of the cursor is expressed as a function of an alteration in the texture. This can produce the illusion of feeling the tactile properties of the textile in question.

Further empirical investigation [3] shows that the C/D ratio is a key component in perceiving the weight of an object. Weight can thereby *"be added to the list of the haptic properties that can be simulated with a pseudo – haptic feedback"*.[3] If the C/D ratio is small, the visual motion of the object in the VE is amplified, thus the object is perceived as being light. On the other hand, the object is perceived as being heavy if the display of user's motion in the VE is lesser than user's control motion. This implies that visual feedback can mislead the cognitive system.

1.2. Presence and Immersion

The term 'presence' refers to the sensation of being included within a mediated reality, and simultaneously accepting this reality as valid. Often defined as the sense of 'being there', users forget their physical surroundings in favour of the virtual world.

Significant to the study of presence, is the criterion by which the experience emerges. A common example of this is how a head mounted display (HMD) immerses the user in a visual representation of the virtual world. Slater and Usoh [5] expand this view to include the virtual representation of the user's body, the virtual body (VB), which must be tracked in the physical world.

> *"Proprioception results in the formation of an unconscious mental model of the person's body and its dynamics. This mental model must match the displayed sensory information concerning the VB."* [5]

Biocca et. al. [1] investigate the relationship between presence and the occurrence of pseudo haptics. The experiments focus on reported haptic feedback as geometrical shapes 'snapping' into the hand of the user. Results confirm a significant correlation between presence and the occurrence of multimodal illusions.

1.3. Objective

The aim of the present investigation is to create a VE where the user registers the haptic property-weight, using only visual cues and a passive input device (no actual haptic feedback device); examining whether this can trigger the illusion of haptics. From this, we investigate the interdependency between the registered sense of presence and pseudo haptic feedback. By applying two different VEs, each with a different level of presence, we hope to uncover the significant influence presence has on pseudo haptic feedback. The problems presented by generating a pseudo haptic feedback among a user audience and analyzing the occurrence of these, are twofold:

- Is it possible to achieve the illusion of haptics through visual cues using a passive input device in a VE?
- Are pseudo - haptic feedback, in the form of a perception of weight, and the sense of presence interrelated?

Based on previous research we hypothesize that an increased sensation of presence will correlate to an increased number of registered pseudo haptics.

2. Technical description

2.1. Test setup

In order to investigate the validity of the hypothesis stated above, two virtual test environments were built with different perceptual load and visual quality.

The experiment was set up in the form of a bowling game where the users competed against each other for a prize. Each of the virtual environments includes bowling balls of three different colours, each with its own weight simulated by the C/D ratio techniques. This was done by making the motion of the virtual hand, holding the bowling ball, slower than the motion of the actual hand for those objects that were to be perceived as heavy. Likewise an acceleration was made in the motion from the actual to the virtual hand, holding the bowling ball, for those that were to be perceived as light.

The player scored a single point for each cone that was knocked over. The intention of the game was to impose an increase in perceptual load on the user by creating a meaningful task. This could contribute to the suspension of disbelief in the users as they would focus their attention on the task rather than on the equipment or the absence of a haptic feedback.

The two environments are illustrated bellow.

Hi Quality - Visual Environment: As shown in Figure 1, the Hi Quality - Visual Environment is set up to encourage the sensation of presence in the user. Through rich textures, a detailed soundscape and interior it is designed to attain the 'atmosphere' of a bowling area.

Figure 1. Hi Quality - Visual Environment

Low Quality - Visual Environment*:* Shown in Figure 2, the Low Quality - Visual Environment is designed to discourage the sensation of presence. This is done by using uniform colours and removing logical elements like the additional interior light sources and the ambient soundscape.

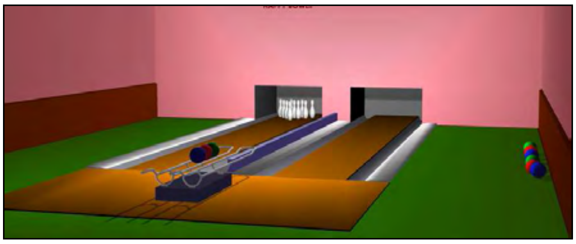


Figure 2. Low Quality - Visual Environment

The technical setup of our experiment consists of various hardware and software.

A Head Mounted Display (HMD) by VRLOGIC with a resolution of 800x500, Polhemus 3Space tracker with two sensors and 6 degrees of freedom, an angular displacement sensor data glove, Teleo microprocessor and computer with two NVIDIA fx1400 graphics card are equipments used in this setup.

The software was used Max/MSP 4.5 by Cycling 74, Virtools 3.5 including physics and VR pack by Dassault Systemes, MIDI Yoke, and VRPN server.

Tracking sensors were placed on top of both the HMD and the data glove tracking the movement with six degrees of freedom. The HMD's control box was connected to two NVIDIA fx1400 graphics card and the Polhemus system electronic unit was connected to the serial port which connects to Virtools using a VRPN server. The data glove connects to the Teleo microprocessor which interfaces with Max/MSP which then sends the angular displacement sensor data to Virtools. A computer screen was also connected in order to see how the users responded to the tasks in the VE.

2.2. Testing

20 undergraduate students participated in the experiment, under the heading 'Virtual Bowling'. The test participants were placed in a location so that they had at least 1.5 meters of free space in all directions. The users were exposed to one of the two virtual environments and given a practise round of 9 bowling balls before the actual game and test started. Here each user again had a total of 9 bowling balls, 3 of each weight, and were instructed to hit as many cones as possible in the virtual bowling game.

Following the test experiment each participant was given a questionnaire of 37 items. The questionnaire is a modified version of Witmer and Singer 'Presence Questionnaire' (PQ) merged with the 'Immersive Tendencies Questionnaire' (ITQ) [6].

2.3. Determinants of the PQ

The PQ was set up to test specific determinants of presence, as created by Witmer and Singer [6]. These determinants are divided into four factors that describe the concept of presence. Each question is rooted in one of these factors.

These factors are further divided into its constructing subscales. The PQ includes a second ITQ which measures the user's willingness to suspend disbelief and submit himself to another reality. The ITQ is also represented in the PQ at the same level as previously defined subscales.

Within this framework, new items could be formulated to suit the specific experimental needs to verify the hypothesis. Three subscales were of specific interest, and were thus assigned more items in the altered questionnarie.

Immediacy of Control is a control factor that emphasizes the user's awareness of delays. The C/D ratio, as used to signify heavier objects, may be perceived as a system showing latency.

The *Active Search* subscale is of significant importance since it relates to the audio/visual sensorial exploration of the environment. Modifications to the PQ extend this measure to the sensation of pseudo - haptic feedback.

Interface Awareness is considered to be a possible distracter for the experiment, since the data-glove has been custom made for the application.

3. Results

A pairwise comparison in the form of a t-test was conducted on the results of the tests performed. This was done to judge the difference between the means of the various questions and factors of the questionnaire, relative to their variance.

In most of the data collected, and measured on an individual basis (30 out of 37 questions), the level of significance was registered to be higher than 0.05; supporting the null hypothesis, that there is not significant difference between the environments.

Subscales were compared between environments and proved not to have a significant difference in all cases but one, the *immediacy of control*.

2.5. On Presence

A between subjects comparison was conducted on reports on presence. The level of reported pseudo - haptics feedback were specifically examined in these comparisons as a direct clue to the validity of the hypothesis. The interest lay is investigating whether there was a correlation between the environments in respect to the reported sensation of presence; also whether there could be registered an interrelation between the level of presence and perception of weight.

Using a five – point Likert scale, the users with a mean average of 3.5 and higher in the presence questions were considered to have registered the sensation of presence. Out of the 20 users, 6 reported a presence mean above 3.5.

Subsequent analysis of the test results related to high and low presence ratings, revealed a correlation between questions relating to the virtual hand. The questions formed a 3.46 mean with a 0.04 variance.

4. Discussion

Taking a closer look at the individual results within that specific subscale, it is evident that the means in both environments are fairly average. Ranging between 3.03 and 2.73 with low and almost equal variances of 0.54 and 0.57; reporting an insignificant diversity on the *Immediacy of Control* factors in the two environments.

A level of significant difference was also registered on individual items. The item on how much the visual display quality distracted the user when playing the game, gave results reporting that users may be less distracted by the visual display in high resolution environments than low resolution.

A total of 6 reported a presence rating above 3.5, equalling 3 participants from each environment. It stands to argue, that a population of 6 subjects in a bi-conditional test environment is insufficient to achieve valid results. Therefore, the subsequent discussion should be understood as estimations.

Users reported as highly present, showed a significant correlation in *Control Factors* across the two environments. A significant determinant in the control mechanism rests in the representational self, by way of the virtual hand. Pertaining to test-subjects with an above average rating of presence, the questions related to the representational self yielded a combined mean of 3.46 with a 0.04 variance.

Users reporting an above average presence also showed a significant ($p < 0.043$) relationship to the occurrence of pseudo - haptic feedback. Items related to the pseudo haptic scale have a mean of 3.42 (high visual resolution) and 2.5 (low visual resolution).

5. Conclusion

As previous research shows [5], self inclusion is an important factor for immersion. By that argument, it is conceivable that the virtual hand added to the sensation of presence for the users in this setup. Furthermore this might underline the significance of using iconic representations of the human body within virtual worlds as opposed to the use of symbolic representations, i.e a 3D curser.

The correlation revealed between presence and the occurrence of pseudo - haptic feedback corresponds to the earlier stated hypotheses: there is a high probability of a correlation between the sensation of presence and the occurrence of pseudo - haptic feedback. The test however, cannot confirm whether or not this correlation is sustainable in a state of growth between presence and pseudo - haptic feedback. What limits this is the low test population, a far higher number, than the estimated 10 pr. condition is required to get a valid reading.

Throughout this experiment, we have uncovered a strong indication, that our hypotheses are correct. The topic of cross modal illusions, remains a dense field and further tests can still be done along the lines of this experiment.

Acknowledgement: we wish to express our gratitude to the studyboard at Aalborg University (AAU), Denmark and to DRF (Advisory Research Committee of the Dean at AAU) for economical support.

References

[1] Biocca F., Inoue Y., Lee A., Polinsky H. and Tang A.. “Visual Cues and Virtual Touch: Role of visual stimuli and intercessory integration in cross-modal haptic illusions and the sense of presence, Meida Interface and Network Design (M.I.N.D) Labs, Michigan State University, 2002

[2] Biocca F., Kim J. and Choi Y., “Visual Touch In Virtual Environments: An Exploratory Study of Presence, Multimodal Interfaces, and Cross-Modal Sensory Illusions”, pp: 247-265 Presence, Vol 10, Issue 3, MIT Press, Cambridge, MA, USA, 2001.

[3] Dominjon L., Lécuyer A., Burkhardt J. M, Richard P. and Richir S. “Influence of Control/Display Ratio on the Perception of Mass of Manipulated Objects in Virtual Environments”, pp: 19-25, Virtual reality, Proceedings, VR 2005, IEEE, 2005

[4] Lécuyer A., Burkhardt J.M. and Etienne L. “Feeling Bumps and Holes without a Haptic Interface: the Perception of Pseudo-Haptic Textures”, Conference on Human Factors in Computing Systems, pp: 239-246, ACM Press, New York, NY, USA, 2004

[5] Slater M. and Usoh M. “Body Centred Interaction in Immersive Virtual Environments”, Department of Computer Science and London Parallel Applications Centre, Queen Mary and Westfield College, University of London, U.K, 1994.

[6] Witmer B. G. and Singer M. J. “Measuring , Presence in Virtual Environments: A Presence Questionnaire”, U.S. Army Research Institute for the Behavioural and Social Sciences, Volume 7, Issue 3, MIT Press, Cambridge, MA, USA, 1998

Designing a interface for virtual reality and haptic application prototyping: a case study

Joan De Boeck [*] Emanuele Ruffaldi [‡] Marcello Carrozzino [‡] Chris Raymaekers [*]
Karin Coninx [*] Massimo Bergamasco [‡]

(*)Hasselt University, Expertise centre for Digital Media and transnationale Universiteit Limburg
Wetenschapspark 2, B-3590 Diepenbeek, Belgium
(‡)PERCRO, Scuola Superiore Sant'Anna, Via Rinaldo Piaggio, 34,
56025 Pontedera, Italy

{joan.deboeck,chris.raymaekers,karin.coninx}@uhasselt.be
{ e.ruffaldi,m.carrozzino,bergamasco}@sssup.it

Abstract

It may sound obvious that computer interfaces strive to adopt the same patterns of learning as we know from our daily world. In particular we refer to the enactive knowledge, acquired by 'doing', such as "driving a car", or "playing an instrument". However, even in highly interactive 3D Virtual Environment applications, the enactive approach is still not common This paper shows how the adoption of a hardware standard, and the use of high level interaction diagrams may support the easier creation of such an interface .

1. Introduction

Technical problems, as well as user interface design problems, may be the main reasons why effective enactive interfaces are difficult to develop. First of all, a rich and multimodal communication channel between human and machine is required for this kind of interfaces, very often resulting in the use of specialized (and hence less standardized) hardware, such as haptic devices and trackers. Moreover, when a specific hardware setup may be operational, the designer has the possibility to choose between dozens of possible interaction techniques to perform a certain task. Some techniques may carry similarities with common techniques in the real world, but they also can act as some kind of "magic" within the virtual world. The most reliable method to know which technique may be suitable in a given situation still consists in testing the proposed setup. This implies that a working prototype must be implemented, which may result in long development times and several iterations of the system Designing an enactive interface clearly is not a simple process. In this paper we therefore describe the design of an application as a case study, in order to prove how hardware standardization and the use of high level interaction diagrams may facilitate the creation of this kind of interfaces.

2. Existing Application

Virtual Prototyping exploits VR technologies to let the users perceive a non physically existing object as real as possible. The aim is to allow the user to perform operations directly on the virtual prototypes with operative modalities almost identical to those used in the real world. The automotive industry is the sector where these technologies have been introduced with the best success, being used to verify complex assemblies of mechanical groups designed by different workgroups[1], to verify the ergonomics of the vehicle interiors, to perform analysis in the recycling stage and to verify the disassembly procedures of the product[2].

PERCRO, in collaboration with Piaggio S.p.A. developed the Interactive 3D Visualization System (IVS) [3], a Virtual Prototyping system

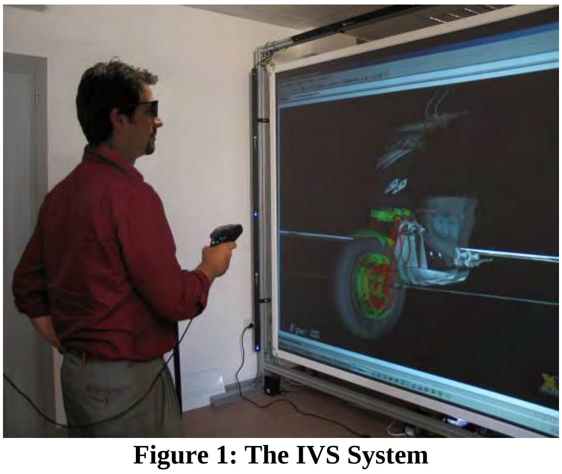


Figure 1: The IVS System

dedicated to the stereoscopic visualization of a 3D model of a motorvehicle in order to verify the final mechanical assembly on a PowerWall-like device. All the software functionalities have been realized with XVR [4], a developing tool for VR applications realized by PERCRO in collaboration with VRMedia.

The system allows to interactively explore the assembly and to manipulate its components. It is possible to select the elements of the assembly and to move them with operations of roto-translation, to temporarily hide them in order to assist the vision of the assembly interior, and to perform measurements of distance between points of the assembly. Two kinds of interaction controllers have been used: a sensorized wand (provided with a joystick and an additional set of buttons associated to various functionalities related to movement, selection etc.) or a glove coupled to a 6-DOF wrist sensor. In the latter case the functionalities are associated to the postures of both hands. In both cases the ray-casting interaction metaphor was chosen for the selection of subcomponents. The two hands are used in different ways: the dominant hand is associated to action tasks (like selecting, moving. etc.) while the non-dominant hand is associated to option tasks (like switching interaction modality, active status etc.). The user's head is also tracked in order to produce a correct dynamic perspective.

Although the overall evaluation was evaluated positively, several issues arose about the interaction metaphors implemented in the system. Although Ray Casting is one of the easiest and most common selection metaphors, it suffers from being less accurate with distant objects because of the relative sensitivity of the rotation of the ray. This behavior discourages users to directly manipulate distant objects, and induce them in performing a preliminary manipulation of the model to bring closer the desired component, in order to allow an easier and more precise manipulation. Another drawback is related to occluded objects which can be accessed only rotating the main model or removing occluding subcomponents.

As the design and evaluation of the initial application was a lengthy process in which each change must be written in code, we have chosen to use this particular application as a case study to evaluate the proposed platforms that may help to facilitate the development and evaluation of an enactive interface.

3. Used Platforms

3.1 VRPN, Virtual Reality Peripheral Network

The haptic rendering framework used for this experimentation is based a new module for the Virtual Reality Peripheral Network protocol. Most of Haptic Rendering libraries are based on a local communication with the rendering library or with simple communication interface with a remote haptic server. Given the required flexibility we have chosen to use the VRPN protocol for providing remote haptic rendering. The contribution of this paper is related to the generalization of the VRPN for multiple contact points and for a flexible approach in the geometry sharing between elements. The use of a remote protocol is extremely useful for also implementing a visual debugging tool of the haptic scene.

3.2 NiMMiT, Notation for Multimodal Interaction Techniques

The creation of (multimodal) user interaction in a 3D environment is a time-consuming and expensive process. Solutions have to be implemented using programming code, and as the acceptance of a multimodal interaction paradigm can not be fully

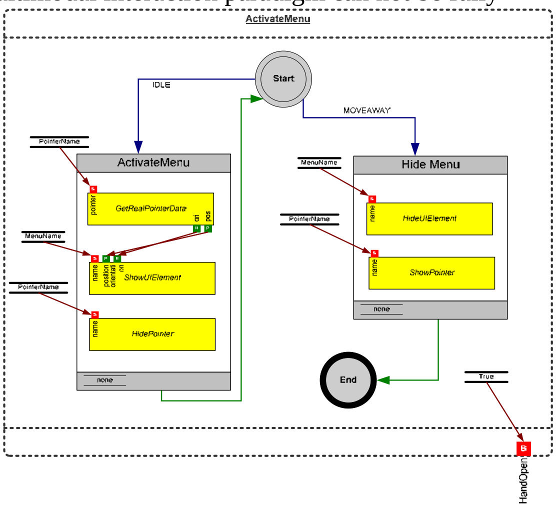


Figure 2: NiMMiT Diagram of the Aperture Selection

predicted in advance, the result often must be evaluated and adapted afterwards. To simplify this process, EDM has developed NiMMiT [6], a graphical notation intended to *describe* multimodal user interaction, rather than *implement* it. The notation can also be used to capture performance data during a user experiment, providing formal data for statistical analysis during the evaluation phase [7]. An example of a NiMMiT diagram is given in Figure 2. NiMMiT inherits the formalism of a state chart, and defines 'events' as a result of the user's actions within the virtual world. In each state, the interaction responds to a given set of events. Dependent on the events occurred, a given set of 'tasks' (grouped in a 'task chain') is executed, executing what is necessary in this particular phase of this interaction technique. For a more comprehensive explanation and more examples of NiMMiT, we refer the interested reader to [6][7][10]. In the remainder of this paper we will explain some simple diagrams, as well.

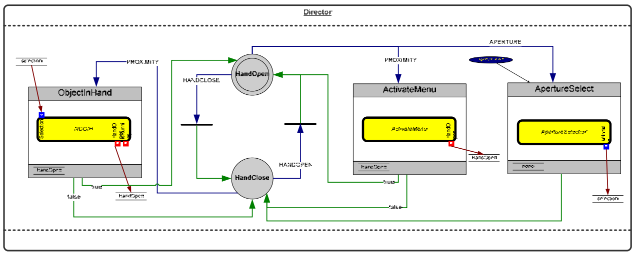


Figure 3: Overall NiMMiT diagram with hierarchical use of the components

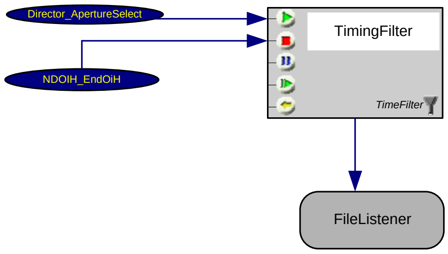


Figure 4: NiMMiT Filter Diagram

4. Case Study: Virtual Prototyping

4.1 Input devices

The proposed experimentation has been prepared with the use of the GRAB haptic interface, a device with two point of contacts that can be used for the experimentation of two hand interaction.

4.2 Interaction Metaphors

4.2.1 Designing Interaction

In the proposed virtual prototyping application, users may have two approaches in order to explore the model. Either they may to explore the entire model, and navigate through and around the model, or they want to 'grasp' certain parts of the model and inspect them separately. The choice between the different inspection techniques can be done using a 3D floating menu.

The menu is activated as described in [8]. By bringing the user's non-dominant hand close to the dominant hand, the menu appears at the position of the user's pointer, so that they can quickly interact with that menu. This approach has proven to be very intuitive, as the user's proprioceptive knowledge is exploited to activate and operate the menu. For the interested reader, we have shown the diagram of this interaction in Figure 3.

For the 'picking' and inspection of individual parts of an assembly, we adopted the 'Object in Hand' metaphor [9], combined with the aperture selection [12], as proposed and evaluated in [11]. Using the non-dominant hand, an 'aperture' is moved onto the screen, highlighting the object that is in the aperture. When the user decides that the desired part is highlighted, this can be 'grasped' by bringing the fist of the non-dominant hand close to the dominant hand. The object then pops out of its context and comes to a central position closer to the user. By moving and rotating the non-dominant hand, the object is affected similarly, allowing the user to inspect it.

As NiMMiT allows hierarchical reuse of diagrams, the combined action of the different parts of the interaction (menu activation, object selection, object grasp and release) can be described by a diagram as well. This is depicted in figure Figure 3. Here we distinguish two separate states, reflecting whether the user's non-dominant hand is open or closed. Dependent on the active state and the event occurring, one of the three task chains is executed. Each task chain contains only one (hierarchical) task, executing the particular NiMMiT diagram.

For the navigation, which can be used for the exploration of the entire model, we have chosen to use the 'scene in hand'-metaphor [13], in which force feedback with the dominant hand, simulating the surface of a sphere using force feedback features, may improve the interaction.

4.2.2 Evaluating Interaction

For the evaluation of the interaction metaphors, we can make use of the Probe and Filter primitives of NiMMiT [7]. For instance, we can easily capture the time that a user needs to select an object and subsequently grasp it. This may be useful, for instance to evaluate whether the combination of the selection technique and the grasping technique is efficient enough. The capturing of the data can be implemented by adding a probe to the 'ApertureSelect' taskchain and another probe at the 'ObjectinHand' taskchain. With the probes, we can receive raw info about the execution of the diagram at these points. Next, in 'filter diagram', the probes can be connected to a 'Filter', processing this raw data. In Figure 4, we can see how the data of the two probes is routed to a timer-filter, using the data of the probes to simply start or stop a timer. The result of the time measurement is then routed to a FileListener, storing the results in a file where it can be used for any statistical analysis afterwards.

5. Discussion

After the time we needed to write an initial interface to VRPN to operate the particular force-feedback device (which only needed to be done once), we could start using VRPN for our project. It turns out that the development of the interface at the labs of EDM, without the availability of the specialized force-feedback devices that are to be used at PERCRO, caused no significant problems. As VRPN acts as a transparent layer between the application and the hardware, only a configuration file at the server -or- at the client side has to be adapted in order to change the input devices.

Designing NiMMiT diagrams is a significant improvement over manually implementing each interaction technique using programming code. First of all, NiMMiT gives the designers an easy to read, easy to understand diagram, allowing them to discuss the proposed solution with other colleagues, when necessary. Next, the key actions in an interaction diagram are the 'tasks'. Dependent on the domain on which NiMMiT is used, several frequently used tasks may be predefined, such as in our case 'moving an object', or 'calculating collision detection'. This speeds up the design a lot, but because designers still have the possibilities to write their own specific tasks ('custom tasks'), NiMMiT's high-level approach does not limit the power we have in low-level solution. In this case study, only one custom task (for the sphere navigation) had to be coded and one task could be reused but had to be adapted from a previous project.

Finally, the addition of features to capture 'real user data' using 'probes' and 'filters' gives the opportunity to quickly add code for interim user experiments, facilitating the iterative aspect of user interface design.

5. Conclusion and Future Work

In this project paper, we described the redesign of an existing Virtual Prototyping application, aiming to improve the enactive aspects of the interface. We focused on two platforms, VRPN and NiMMiT, both facilitating the development process. VRPN makes abstraction of the concrete hardware device, avoiding the need to have specific and expensive hardware to be available at all times. NiMMiT on its turn allows a designer to easily describe the user interaction and adds additional features to capture real user performance data for an evaluation of the designed interface.

We can conclude that our approach improves the flexibility of Interface Design, and especially addresses the experimental approach of creating enactive interfaces by facilitating the iterative process of choosing devices, designing interaction techniques, evaluating the result and iterating over again. Taking into account the possible reuse of previous investments, we can state that this strategy allows an interface designer to deliver an interface prototype in much less time than by coding it manually.

Concerning the case study described in this paper, after extending the features of the application we want to conduct a formal usability experiment, evaluating the different interaction techniques. As NiMMiT is currently still running on top of an experimental research framework (VRment), it may be valuable to integrate it in a more powerful environment such as XVR, as well.

6. Acknowledgements

Part of the research at EDM is funded by EFRO (European Fund for Regional Development), the Flemish Government and the Flemish Interdisciplinary institute for Broadband Technology (IBBT).

Part of the research at PERCRO and EDM has been funded by European Community in the Enactive NoE

Reference

[1] Niesen M. R. and Luecke G. R. "Virtual dynamic protptyping for operator interface design". Proc. Of IEEE ro-Man, September 27-28 1999.

[2] Roy U, Bharadwaj B, Kodkani SS, Cargain M. "Product development in a collaborative design environment." Concurrent Engineering Research And Applications 1997

[3] Marcello Carrozzino, Andrea Baroni , Franco Tecchia,Fabio Rossi,Massimo Bergamasco :"Interacting with complex virtual models: a suitable metaphor for Virtual Prototyping", proceedings of Enactive Conference 2005

[4] http:www.vrmedia.eu

[5] Tailor II, R. M. (August 2006). VRPN virtual reality peripheral network. http://www.cs.unc.edu/Research/vrpn/index.html.

[6] Vanacken, D., De Boeck, J., Raymaekers, C., and Coninx, K. (2006). NiMMiT: A notation for modeling multimodal interaction techniques. In Proceedings of the International Conference on Computer Graphics Theory and Applications (GRAPP06), Setbal, Portugal.

[7] Coninx, K., Cuppens, E., De Boeck, J., and Raymaekers, C. (2006). Integrating support for usability evaluation into high level interaction descriptions with NiMMiT. In Proceedings of 13th International Workshop on Design, Specification and Verification of Interactive Systems (DSVIS'06), Dublin, Ireland.

[8] De Boeck, J., Raymaekers, C., and Coninx, K. (2006d). Exploiting proprioception to improve haptic interaction in a virtual environment. In Presence: Teleoperators and Virtual Environments, volume 16. The MIT Press.

[9] De Boeck, J., Cuppens, E., De Weyer, T., Raymaekers, C., and Coninx, K. (2004). Multisensory interaction metaphors with haptics and proprioception in virtual environments. In Proceedings of the third ACM Nordic Conference on Human-Computer Interaction (NordiCHI 2004), Tampere, FI.

[10] De Boeck, J., Raymaekers, C., and Coninx, K. (2006c). Comparing NiMMiT and data-driven notations for describing multimodal interaction. In Tamodia 2006, Diepenbeek, Belgium.

[11] De Boeck, J. A User and Designer Perspective on Multimodal Interaction in 3D Environments, PhD Thesis, Diepenbeek, 2007

[12] Forsberg, A., Herndon, K., and Zeleznik, R. (1996). Aperture based selection for immersive virtual environments. In Proceedings of UIST96, pages 95–96.

[13] Ware, C. and Osborne, S. (1990). Exploration and virtual camera control in virtual three dimentional environments. In Computer Graphics, volume 24, 1990.

The Speed Accuracy Trade–Off through tuning tasks

Amalia de Götzen* Davide Rocchesso†

(*) Sound and Music Computing group, University of Padova, Italy, http://smc.dei.unipd.it
(†) Faculty and Arts and Design, University of Venezia, Italy
E-mail: degotzen@dei.unipd.it, roc@iuav.it

Abstract

This paper present the results of an experiment that deal with sound perception, human performance, audio processing and interaction design. The fil rouge that weaves all these topics together is the use of a continuous audio feedback in designing interfaces in order to better understand human gestures and feedbackgesture coordination. The aim of the experiment is to transpose in the auditory domain predictive HCI models in order to evaluate how humans perform having audio feedback as the main source of information. In particular this work focuses on the discussion of tuning tasks experiments, using Schmidts law as predictive model. A better accuracy at faster speeds has been observed: the plots characterized by a slower speed show a better but steeper linear dependence between the amplitude of the interval and the standard deviation, no matter what the frequency range of the interval is. Thus, accuracy with auditory feedback seems to observe different dependencies from speed than those predicted by Schmidts law having just the visual feedback allowing the authors for further investigation.

1. Introduction

The importance of predictive models in Human–Computer Interaction is due to the possibility of evaluating human performance, without actually realizing expensive and timeconsuming experiments. These models can guide designers in the choice of an interaction strategy, or they can help in choosing the most efficient input device before observing it in use. This is especially useful when there are basic gestures that recur often and may represent a bottleneck in using an interface. In the age of "gestural interfaces" users are very often asked to iterate basic gestures and movements to play an engaging game or to control an appliance.

In 1954 Paul Fitts published the first paper [1] about a mathematical model applied to the human motor system. He pointed out that the human target acquisition performance could be measured in information units using some indexes derived from the Shannon Theory of Information [4]. Around 20 years after Human Computer Interaction researchers deeply investigated the use of Fitts' law as a predictive model to estimate the accomplishment time of a given task, or to compare different input devices. Nowadays, Fitts' law is coded in an ISO standard [2] and it is still under debate for several aspects: the mathematical formulation, the theoretical derivation, the range of applications, etc. The model is based on time and distance. It enables the prediction of human movement and human motion based on rapid, aimed movement. Fitts discovered that movement time was a logarithmic function of distance when target size was held constant, and that movement time was also a logarithmic function of target size when distance was held constant. Mathematically, the ISO standard version of Fitts' law is stated as follows:

$$MT = a + b\log_2(A/W + 1), \qquad (1)$$

- MT = movement time
- a,b = regression coefficient
- A = distance of movement from start to target
- W = width of the target

Schmidt's law [3] is often studied in comparison with the Fitts' law. The main difference between the two approaches is that Fitts treated movement amplitude (A) and target width (W) as independent variables and movement time (MT) as a dependent variable, while Schmidt chose to manipulate amplitude and movement time, and to measure the movement variability of the effective target. In the Schmidt's case the impulse variability model forecasts that the standard deviation in end-point coordinates (viz., accuracy), is a linear function of velocity, calculated as distance over time:

$$W_e = a + bA/MT \qquad (2)$$

It is worth noticing that the previous equation and equation 1 contain the same three parameters (except that We is the standard deviation of end-point coordinates and it amounts to $4.133 \times SD$ in Fitts' adjusted model). Although this equation can be rearranged with MT as the predicted variable, it is still fundamentally different from the Fitts' law since the relationship is linear rather than logarithmic, and because the information–theoretic analogy is absent. It is interesting for our research to investigate this law with audio feedback since temporal constraints are very often present when we are dealing with music and performance tasks. When a music performer reads a score he has to follow a specific timing (that can be varied only slightly by his own interpretation). This means that he has to perform a Schmidt's task: reaching a target within a timing (MT) and an interval (A) (as in equation 2). The metaphor that inspired our experiment is that of a glissando made by a violin player: following a certain tempo he has to reach a given position on the fingerboard (a typical Fitts'/Schmidt's task, with spatial constraint), but she has also to reach a given note which corresponds to that specific position. The feedback that a musician uses performing this kind of tasks is both kinesthetic and auditory.

2 The experiment

The idea of the experiment is to explore the speed/accuracy trade off when there are temporal constraints in order to discover if the Schmidt's law is still reliable with auditory feedback. Twenty–one subjects (between 21 and 48 years old) participated in the experiment performing 24 trials each. All participants reported normal hearing and sight, and normal motor capabilities in their hands. All of them were naive as to the purpose and hypotheses of the test and all of them volunteered. The test has been carried out using Pure Data software.

Auditory stimuli were obtained using simple pure tones: the user has to reach a target from a starting frequency. The 8 different tasks which corresponded to 8 different intervals had to be performed at 3 different linear speeds (time constraints). The intervals have been chosen in order to be non-musical intervals and to explore the whole audible spectrum. They were proposed in random order and subjects could practice with the tasks before starting the test (the practicing data set was dropped for every subject). The test was not intended to measure the reaction time of the subjects, but the data dispersion, measuring the distance between the real and the performed target. It attempted to mimic closely the visual test, where the target and the distance remain visible throughout each task, therefore, at the beginning of each trial, the starting frequency, the ending frequency, and the glissando between the two were heard. After this initial phase the subject actually performed the trial. No visual feedback was provided. She/he was simply supposed to press a button in order to start and stop the controlled frequency while also continuously hearing the target. Table 1 shows the list of intervals that have been used for the three different linear speeds (fast 1Hz/ms, medium 0.2 Hz/ms, slow 0.111 Hz/ms):

Trial	Initial freq. (Hz)	Final freq (Hz)
t1	340	1310
t2	340	2430
t3	340	3270
t4	1240	2510
t5	1240	3700
t6	2170	3580
t7	2170	4000
t8	3300	4020

Table 1: Initial and target frequency of each trial

2.1 Data analysis

A one factor ANOVA test showed the significance of the mean value of the collected data: the computed p factor was definitely below the significance threshold (p=0.05). Fig. 1 shows several aspects at the same time: the blue bars in the upper plot describe the distances (A) in Hz between the initial frequency and the target one, while the red crosses represent the mean values of the reached targets for each trial. The lower plot shows the corresponding mean deviations from the target frequencies. The width of each bar represents the distance between the starting frequency and the ending frequency. Thus a wider interval will correspond to a wider bar and vice-versa. Even though each subject shows her/his own characteristic behavior some common patterns may be noticed: the task seems to be easier when the speed is higher.

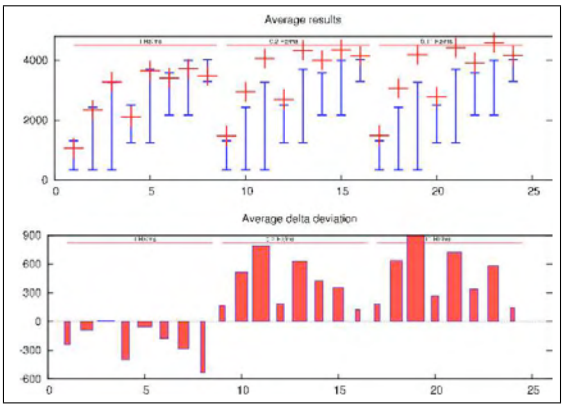


Figure 1: Reached frequency Mean values. In both plots the unit of the vertical axisis Hz, while that of the horizontal axis is the number of the task.

As a matter of fact fact fig. 1 clearly shows that the mean deviation is bigger for lower speeds, where participants often stop the glissando before reaching the target frequency. Summarizing:

- the performances indicate two different behaviors; at high speed subjects tend to stop the glissando before reaching the target (temporal constraint), while they tend to stop it *after* reaching the target when the performance is at slow speeds;
- the standard deviation is decreasing with increasing speeds but it will be shown that the results have better uniformity for the 8 trials at the slowest speed;

Figures 2 3 show the boxplots of the data callecetd at the highest and slowest speed: the horizontal bar in the middle of each box shows the median value of the reached frequency. The top of the box above the median shows the third quartile, and the bottom of the box below the median shows the second quartile. The two boxes together show where the middle 50% of the data lie, while the whiskers show the maximum and minimum values of the reached frequency. The outliers are the small circles outside the boxes. Figure 2 shows the boxplots of the data collected at speed 1Hz/ms. Even if there are some outliers, a smaller dispersion of the data can be observed in comparison with figure 3. Assuming that we can treat the speed of played intervals as the speed of time-constraint movements, audio feedback may follow Schmidt's law: the data regression lines are plotted in figures 4, 5.

Figures. 2, 3, 4, 5 show that the relationship between W_e and A/MT is verified for every time constraint, though the fit is uneven at the highest speed. Schmidt's law states that the relationship between W_e and A/MT is nearly linear: for different amplitudes and movement times the W_e is reported to be about the same when the ratio A/MT is kept constant. Figures 4 and 5 indicate the W_e trend for each A/MT: each point is specified by the amplitude A (the distance between the initial frequency and the target frequency) and the standard deviation (W_e); a regression line is then computed.

2.1.1 Results

The linear speed/accuracy trade-off was observed by Schmidt in very rapid actions where there was probably not enough time to detect errors and issue a correction. The linear trade-off was originally observed using controlled MTs (longer than the shortest MT), whereas in the Fitts' paradigm the goal MT was to be as fast as possible while maintaining a high accuracy. Both these constraints appear to influence the behavior. Regarding this question, Schmidt himself asserts that:

- a linear trade-off seems to occur for movement tasks that are pre-programmed - under open loop control

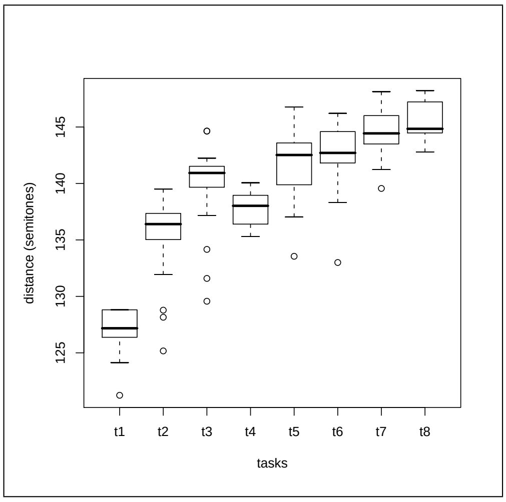


Figure 2: Data collected at 1 Hz/ms (semitones): boxplot

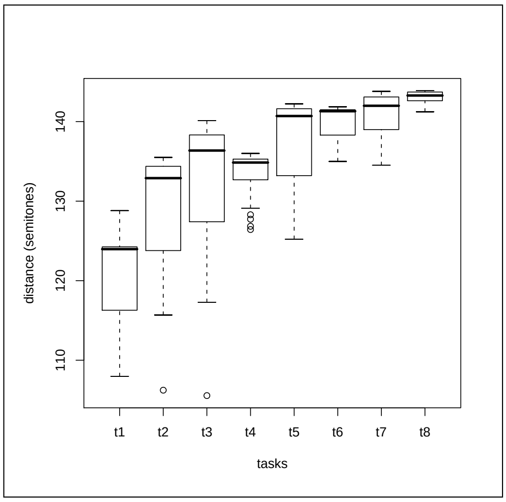


Figure 3: Data collected at 0.111 Hz/ms (semitones): boxplot

- a logarithmic trade-off seems to occur for movement tasks that are governed by feedback-based corrections - under closed loop control

So, while the question of whether the relation is linear or logarithmic may seem rather abstruse, it has a functional significance if a movement is carried out in an open loop or is subject to feedback-based corrections. The difference with the original Schmidt's task is that in our case there is no way to correct the answer even when the movement is slower, since the glissando can be just stopped and not moved back or forth. Thus, this task is always an open loop one and a linear trade off is expected in all cases.

The comparison between the three regression plots in logarithmic scale indicates that the the effective target in the audio domain seems to be, indeed, strictly propor-

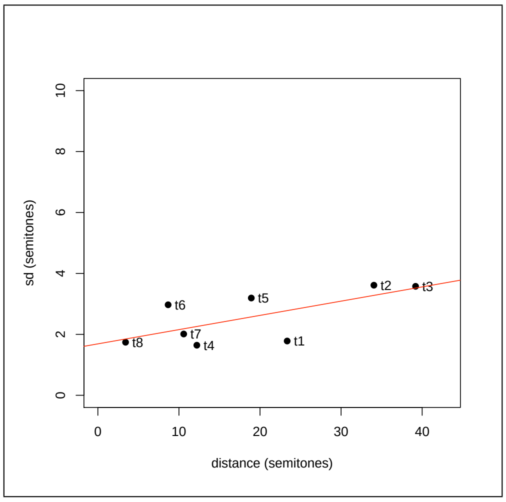


Figure 4: Data collected at 1 Hz/ms (semitones): regression line

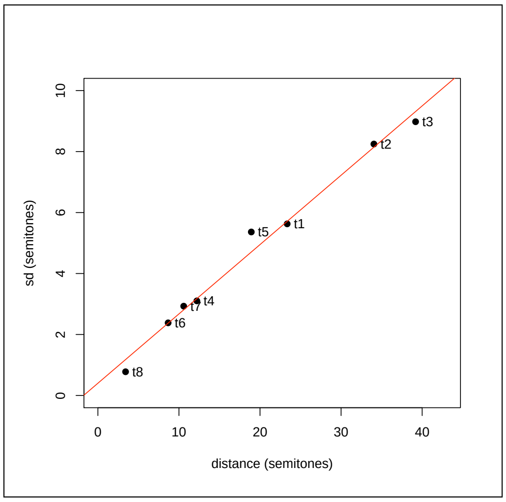


Figure 5: Data collected at 0.111 Hz/ms (semitones)

tional to the distance A in all cases, while the frequency range in which our "gesture" is located does not influence the performance precision. While this linear behavior is exactly what is expected from Schmidt's law and can be derived from the formula 2, the better accuracy at faster speeds observed by looking at the inclination of the regression lines or at the size of the boxplots proposes a different perspective than Schmidt's experiments. The plots characterized by a slower speed show a better but steeper linear dependence between the amplitude of the interval and the standard deviation, no matter what the frequency range of the interval is. One last note on the data collected at the fastest speed (1 Hz/ms) is in order: the regression plot at this speed shows a worse fitting of a linear dependence between distance and standard deviation, even though the smaller standard deviations show better precision. This could indicate that the speed used in this test hits some psycho/physio–logical barrier of another kind: the results maybe biased, for example, by the subject's reaction time.

The findings of this experiment clearly suggest the need of deeper investigation on predictive laws in HCI with audio feedback: many variables and components of audio perception are at work in this kind of experiments, and it is necessary to analyze them and their effects on user performance.

3 Conclusions

The data analysis has shown that the effective target in the audio domain seems to be, indeed, strictly proportional to the distance A in all cases, while the frequency range in which our "gesture" is located does not influence the performance precision. While this linear behavior is exactly what is expected from Schmidt's law and can be derived from the formula 2, the better accuracy at faster speeds observed by looking at the inclination of the regression lines or at the size of the boxplots proposes a different perspective than the original Schmidt's experiments. The plots characterized by a slower speed show a better but steeper linear dependence between the amplitude of the interval and the standard deviation, no matter what the frequency range of the interval is. Thus, accuracy with auditory feedback seems to observe different dependencies from speed than those predicted by Schmidt's law. This is indeed a particular achievement of the Schmidt's–like test: while it is predicted that a smaller speed should allow the subject an easier hit of the target frequency, what happens is exactly the opposite: tests show better results with higher speeds, that is when the target should be more difficult to be hit.

References

[1] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, Vol. 47(No.6), June 1954.

[2] Ergonomic requirements for office work with visual display terminals (vdts)–part 9–requirements for non-keyboard input devices (iso 9241-9) international organisation for standardisation. *ISO 9241-9:2000(E)*, 15, 2002.

[3] R. Schmidt, A. Zelaznik, B. Hawkins, J. Frank, and J. Quinn. Motor output variability: A theory for the accuracy of rapid motor acts. *Psychological Review*, 86:415–451, 1979.

[4] C. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, pages pp. 379–423, 623–656, July-October 1948.

Exploring similarities of affective and sensorial expressive intentions in music performance

Giovanni De Poli   Luca Mion   Ennio Rapanà
CSC-DEI, University of Padova, Italy
E-mail: giovanni.depoli@dei.unipd.it, luca.mion@dei.unipd.it, enniorapana@yahoo.it

Abstract

Physical and perceptual similarities of affective and sensorial expressive intentions in music performance are explored. Machine learning techniques were employed to select and validate the most relevant low level features and an interpretation of the clustered organization based on action and physical analogy is proposed. A perceptual experiment is then presented which confirms the same groupings and suggests an intrinsic correspondence of affective and sensorial expressive intentions.

1. Introduction

Musicians enrich their performances with expressive intentions acting on their available freedom degrees [1], and introducing deviations from a mechanical playing of the score. Most studies of performance expressiveness understanding analyse these systematic deviations [2] and develop generative models for musical applications [3]. In [4] we addressed the question whether expressive information can be communicated (and recognized) by means of features which are not strictly related to the score. Thus, relevant musical attributes for differentiating expressions (such as articulation) can be replaced by more physical features (e.g. the attack time). By machine learning techniques we found the most relevant audio features allowing to recognize different expressive intentions.

In this work we investigate the expressive intention similarities in the feature space and we propose an interpretation of the resulting clustering based on action metaphor and on ideal physical systems. In this way we approach the description of music expression from an intermediate level, between music intended as a structured language, and sound as audio waveform, leading to more robust description of expression and less dependent on different cultures and musical epochs. Moreover the action metaphor could be extended to (and shared with) other kind of dynamic and gesture based arts, such as dance, drawing etc. [5].

2. Expressive music performances

Most research on music expression deals with emotions. However, expressive intention is a broader concept that comprehends emotions, but also sensorial or metaphoric aspects. We considered expressions that were extensively used in previous researches on music expression, and that were confirmed in many experiments to be robust. In the affective domain, emotions are consistently represented on a two dimensional space called the *Valence-Arousal* space (affective space). In the sensorial domain we considered Kinematics-Energy space [6], which was derived analysing expressive intentions inspired by sensorial adjectivesIn this space the first dimension is characterized by bright-light vs. heavy performances, the second one by soft vs. hard performances. The first dimension is closely correlated with Tempo and can be interpreted as *Kinematics* factor, while second one is related to attack time, Legato/Staccato, Intensity and can be interpreted as *Energy* factor. Normally affective and sensorial domains are studied separately. In our experiment we want to study the expressive content conveyed by the performers from a more general point of view. Thus, we took into account both spaces (Fig. 1) by using two pairs of opposite labels to indicate the dimensions for each space. Regarding the affective space, the categories Happy-Sad (High and Low Valence), Angry-Calm (High and Low Arousal) represent the bipolarity induced by indepen-

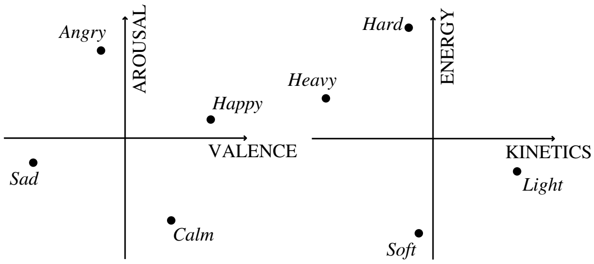


Figure 1: The Valence-Arousal space (left) and the Kinematics-Energy space (right), respectively.

dent dimensions valence and arousal; for the sensorial space, we have the correspondences Hard-Soft (High and Low energy) and Light-Heavy (High and Low Kinematics). In this way each adjective has its opposite in order to deliberately induce contrasting performances by the musician. Beyond the pair of adjectives representing the bipolarities of the spaces, we considered a neutral performance as well, which listeners placed between the pair of opposite adjectives [6]. By "neutral" we intend a human performance without any specific expressive intention and stylistic choice.

3. Representing expression in the feature space

In order to represent similarities of expressions we used machine learning techniques to select the most relevant audio features allowing to recognize different expressive intentions both in the affective (Valence-Arousal) and sensorial (Kinematics-Energy) spaces [4]. First we extracted many audio features from a set of expressive performances played by professional musicians on various instruments. Then we applied Sequential Forward Selection (SFS) with reference to a Minimum Distance classifier to rank and select a set of relevant features. By Principal Component Analysis (PCA) on the performance data we removed correlated features and projected on a 2D space. As a result, we derived a set of features for a general description of the expressions and another one specific for each instrument. These features were tested and confirmed by the leave-one-out cross validation.

Relevant features: As *event* features we found: Peak Sound Level $PSL = \max[RMS(t)]$, where $RMS(t)$ is the temporal envelope; Attack time A as the time required to reach the $RMS(t)$ peak, starting from the onset instant; Notes per Second NPS which is computed by dividing the number of onsets by the window duration. For computing event features we segmented the signal by onset detection, based both on the derivative of the spectral magnitude and on pitch-tracking approach. The offset instant was detected when the temporal envelope $RMS(t)$ falls by the 60% from its previous maximal value.

As *local* features we found relevant: Roughness R which is considered to be a sensorial process highly related to sound texture perception and is computed as in [7]; Spectral Ratio SR_a, which indicate the relative amount of energy in the low frequency band LB ($f < 1$ kHz) and is computed by $SR_a = \sum_{j \in LB} |X(j)|^2 / E_x$, where $E_x = \sum_{k=1}^{N/2-1} |X(k)|^2$is the signal energy; Residual Energy ratio RE_h, which describes the stochastic energy in the high frequency band HB ($f > 1.8$ kHz), obtained by removing the sinusoidal components, and gives information on the quality of the perceived effort. RE_h can be computed by $RE_h = \sum_{j \in HB} |X_R(j)|^2 / E_x$, where X_R is the spectrum of the residual component.

Similarities in the feature space. To understand how the expressive performances, represented by the selected features, are projected (and clustered) on a low-dimensional space, we applied PCA analysis and we used the k-means algorithm for unsupervised clustering of performances from instruments (separately) with the to facilitate the cluster-based interpretation of our features. This iterative partitioning minimizes the sum, over all clusters, of the within-cluster sums of point-to-cluster-centroid distances. We used cosine distance metrics instead of the default squared Euclidean distances employed by k-means: the cosine distance metric compares the angle between two signals rather than the difference in magnitude which is what many other criteria distance measures. This property of the cosine distance metric helps make it invariant against relative scaling of the features. In particular, cosine distance of two feature vector $\mathbf{x}$ and $\mathbf{y}$ is computed by $d = 1 - \cos(\mathbf{x}, \mathbf{y}) = 1 - <\mathbf{x}, \mathbf{y}> / ||\mathbf{x}|| * ||\mathbf{y}||$. In Fig. 2 we show the PCA

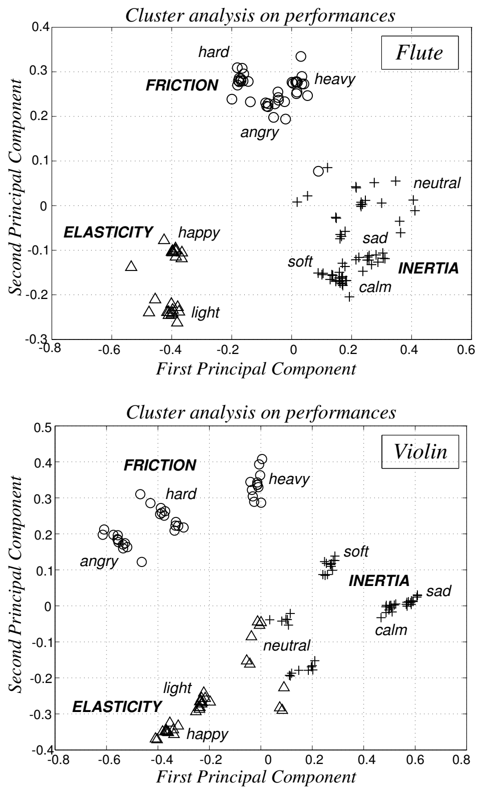


Figure 2: PCA projection and cluster analysis of flute (top) and violin (bottom) performances

projection on a joint 2D space and cluster analysis of flute and violin performances separately.

Interpretation of the similarities. We can exploit the significance of the selected features in order to see how all the adjectives are jointly organized in the feature space, and to explain the found relations relating the space to a semantic interpretation and possible association among sensorial and affective labels. We can see in Fig. 2 that three main clusters emerge: (a) Hard/Heavy/Angry, (b) Sad/Calm/Soft, (c) Light/Happy. An interpretation of the clusters based on action and physical analogy can be devised. When using the physical analogy, force is often subjectively considered as the cause and movement as the effect. The cause-effect relation is represented by the admittance Y which mathematically describes the dynamic mapping and the qualitative behaviour from force to velocity by an integral-differential equation. We can distinguish resistive admittance, which dissipates energy, from reactive impedance, which stores energy. In linear mechanical systems three elementary relations define the fundamental quantities friction, inertia and elasticity. Ideal friction is a pure resistive admittance, while ideal inertia and elasticity are pure reactive admittances: in particular inertia stores kinematics energy and it opposes changes in movement, while elasticity stores potential energy and opposes changes in forces. In general the admittance is composed by a resistive part and an reactive part. In order to have an intuitive idea of their behavior, in Fig. 3) we represent the output (velocity v(t)) from the three basic elements when a smoothed large force pulse is presented at their input. It can be seen that friction acts as a scaling factor of the input force and does not modify the shape of the input. The inertia (mass) tends to remain at its initial velocity, which is zero in the present example. Then it grows progressively and remains constant when the input stops; the mass progressively augments its kinetics energy. The elasticity (spring) instead reacts immediately to the input variations; it stores potential energy which is used to opposes to force changes. From this qualitative description of the behavior of the three basic elements we are induced to associate *friction* to the cluster Hard/Heavy/Angry, *inertia* to cluster Sad/Calm/Soft, and *elasticity* to cluster Light/Happy.

4. Subjective similarities

In order to investigate subjective similarities, we designed an experiment to study how the listeners associate the 9 expressive intentions (4 emotional, 4 sensorial plus neutral) to 3 performances representative of our clusters (friction, elasticity and inertia). A total of 16 participants (musicians and not musicians) participated to the experiment.

Material. Two professional performers of violin and flute were asked to play musical performances in order to convey different expressive intentions, represented by the adjectives happy, sad, angry, and calm (affective space), light, heavy, soft, and hard (sensorial space) plus a neutral performance. Performances are selected from

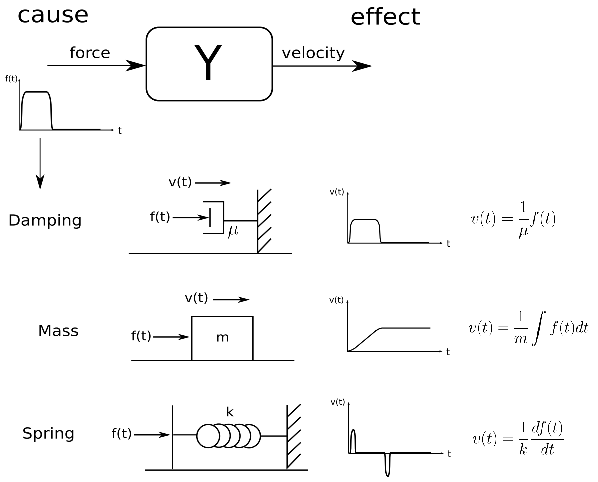


Figure 3: Behavior of the basic linear mechanical systems: friction, inertia, elasticity

the traditional and classical repertoire: Twinkle Twinkle Little Star and a theme from Handel's Sonata in E minor. In total we collected 2 x 2 x 9 = 36 examples, resulting in 4 independent trials for each test. In addition, we recorded 9 scales played by violin (with extreme expressive emphasis) in order to convey the same expressive intentions; for each group of expressions belonging to a cluster, we selected the central one (minimizing the cosine distance) in order to represent the related cluster. Thus, we selected a scale played in order to convey *hard* to represent the friction cluster, *happy* for the elasticity cluster and *calm* for the inertia cluster.

Apparatus. The sound files are presented on the computer screen by a visual interface with 36 buttons (one for each performance presented in random order) and in addition 3 buttons to re-listen to the three example scales at any time. Each of the 36 buttons is associated to a radio button where the participants can only select one choice (a, b, or c).

Procedure. At the beginning of each test, participant is invited to listen to the scales as many times as he/she wishes as training phase. Then, participant is asked to listen to the 36 excerpts and, for each excerpt, to associate the expressive intention that performer wants to convey selecting the corresponding scale on the radio button. Participant is allowed to listen to the excerpts, to the scales and to reselect the corresponding scales as many time as wished. Also, the participant is allowed to stop the performance playback at any time if ready to make the association choice. Each test lasts about 15 minutes. After the test, participant is invited to describe with one or more adjective the expressive content that performer wants to convey with each scale.

Results. The first goal of the analysis was to establish if the preferences in associating the adjectives to the three clusters are accidental or if there is a significant influence on the frequency distribution. Pearson's χ^2 test was used to compare observed frequencies with expected frequenciesfor the total responses and for expressions separately. We obtained a two-way contingency table for the adjectives (9 statistical units) and clusters (3 modalities). Chi-square test yielded to a total value of $\chi^2 = 401.99$ greater than the critical value (36.12 with a level of significance $\alpha = 0.001$, degree of freedom $df = 14$) demonstrating that there is a strong dependence between clusters and adjectives. The values of χ^2 for each expression led to the same conclusion except for the neutral expression. Thus we can argue that listeners did not categorize neutral performance, as expected by the performer.

A second issue was to graphically represent the strength of associations between adjectives and clusters through the technique of Simple Correspondence Analysis on the contingency table. This technique aims at determine scores describing how similar or different responses from two or more variables are (not considering at this point the neutral expression). We obtained a two-dimensional solution for each instrument. In particular, the expressions for each cluster giving the greatest contribution to the total moment of inertia are angry, happy and sad (with an average amount of moment of inertia 0.068, 0.19 and 0.176 respectively), In Fig. 4 we can also notice a distribution of points shaped as an arch (Guttman effect) suggesting a mono-dimensional explanation of this phenomenon, we are further investigating. The k-means algorithm was used to cluster expressions trying to minimize at each iteration the within-group inertia. Three stable groups were identified: (a) Hard/Heavy/Angry, (b) Sad/Calm/Soft/, (c) Light/Happy, with the expressions hard, happy, calm near to the centroid of respective cluster These results are in perfect agreement with previous k-means clustering on the feature space (see Fig. 2) and may indicate an intrinsic similarity of affective and sensorial expressive intentions.

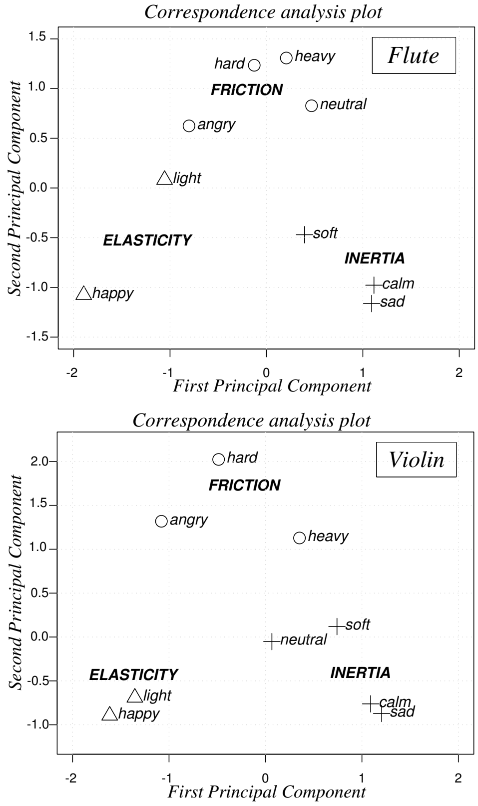


Figure 4: Correspondence plots.

5. Conclusions

An action-based interpretation for expression in music performances is proposed, aiming at a more robust description for music expression, and independent of different cultures and musical epochs. Perceptual experiment showed a perfect agreement between the responses and the clustering on the feature space, suggesting an intrinsic correspondence of affective and sensorial expressive intentions. Suggested physical analogy may be extended to other dynamic and gesture based arts, such as dance and drawing.

References

[1] G. De Poli. Methodologies for expressiveness modelling of and for music performance. *J. New Music Research*, 33(3): 189–202, 2004.

[2] A. Gabrielsson. The performance of music. In D. Deutsch, editor, *The psychology of music (2nd ed.)*, pp. 501–602. Academic Press, San Diego, CA, 1999.

[3] G. Widmer and W. Goebl. Computational models of expressive music performance: the state of the art. *J. New Music Research*, 33: 211–224, 2004.

[4] L. Mion and G. De Poli. Score-independent audio features for description of music expression. *IEEE Trans. Speech, Audio and Language Process.*, 15, 2007.

[5] G. De Poli, et al. Towards a multi-layer architecture for multi-modal rendering of expressive actions, In *Proc. 2nd Int. Conf. on Enactive Interfaces*, 2005.

[6] S. Canazza, G. De Poli, A. Rodà, and A. Vidolin. An abstract control space for communication of sensory expressive intentions in music performance. *J. New Music Research*, 32(3): 281–294, 2003.

[7] M. Leman. Visualization and calculation of roughness of acoustical musical signals using the synchronization index model (SIM). In *Proc. DAFX-00*, p. 125–130, 2000.

Empirical Evaluation Assistant Tool for 3D Interaction Techniques

Christophe Domingues* Samir Otmane* Frédéric Davesne* Malik Mallem*
(*) IBISC Laboratory, University of Evry/CNRS FRE 2873, France.
E-mail: christophe.domingues@ibisc.univ-evry.fr, samir.otmane@ibisc.univ-evry.fr, frederic.davesne@ibisc.univ-evry.fr, malik.mallem@ibisc.univ-evry.fr

Abstract

Designing usable and effective 3D Interaction Technique (3D IT) is very challenging for system developers and human factors specialists. Indeed, time consuming empirical evaluation is necessary to have an idea about the goodness of 3D IT at the end of its development lifecycle. This may induce a huge loss of time if the result appears to be under expectations at the end.

We have developed an Empirical Evaluation Assistant (EEA) to rapidly gather significant feedbacks about the usability of 3D IT during its development lifecycle. Thus, it may be possible to enhance iteratively the 3D IT before it would be classically evaluated by ergonomics experts at the end of its development lifecycle.

EEA has been used to gather feedbacks about a 3D IT developed at IBISC laboratory, called Follow-Me, which is still under study. Results show that EEA has permitted to refine some characteristics of this 3D IT.

Keyword: Human computer interaction; 3D Interaction; Empirical evaluation; Usability.

1. Introduction

At IBISC laboratory, we have been creating interaction models and techniques for our semi-immersive Virtual Reality/Augmented Reality EVR@ platform, especially dedicated to robot teleoperation and collaborative telework.

The most important constraint we are facing is the usability of our techniques. However, there exists no true guidelines to fulfill this constraint when building and implementing 3D IT on a VR/AR platform. Thus, the only choice we had was to validate our IT by ergonomics experts at the end of their development lifecycle. But this validating phase takes a long time and if it appears at the end that the result is poor, the validation feedback comes too late.

Our idea is to build an Empirical Evaluation Assistant (EEA) in order to perform *light* evaluations of 3D IT during its development lifecycle, without the need of being an ergonomics expert. We want this tool to:

- Bring assistance to fasten the preparation of the validating experiments by using information (pertinent variable, experimental protocols, etc) collected in the *knowledge database*;
- Bring fast feedbacks about a tested 3D IT during the validation experiments;
- Collect data and enrich the *knowledge database* to increase the laboratory knowledge about 3D IT behaviors, after the validation experiments.

We hope the EEA system permits to improve the quality of our 3D IT (*Follow-Me*), leading in most cases to positive final evaluations by ergonomics experts.

This paper is structured as follows. Section 2 will briefly review the classical kinds of ergonomic evaluations. The EEA system is developed in section 3. Section 4 gives some insights about the feedback given by EEA with the *Follow-Me* 3D IT.

2. Related work

Ergonomic evaluation is a mandatory step to detect usability problems for creating intuitive and transparent interactions for users. Interaction Techniques for VE are completely different from classical IT for classical Human Computer Interaction. Classical IT is typically used with a keyboard and a mouse to manipulate graphical interface (WIMP paradigm). Whereas their exist guidelines based on predictive models to build effective classical IT, it is not the case for IT for VE. The main reasons are: no strong models, new interfaces and devices, fewer experts. Indeed, two kinds of evaluation approaches exist for IT developed for VE:

- The *analytical approach* compares the behavior of the interaction to a reference model, which describes the conditions to obtain a good interaction (heuristic evaluation, summative evaluation). There are tools dedicated to this approach like the MAUVE system [1], which provides a structured approach for achieving usability in VE system design and evaluation.
- The *empirical approach* measures the performance of different users that are using the IT in

the VE [2]. Due to a lack of norms and ergonomics experience, analytical approaches cannot be used to evaluate [2].

Indeed, empirical evaluation needs to be carried out for evaluating 3D IT (see [2], [5] and [6]). However empirical evaluations are complex to perform due to main difficulties: large list of parameters like users' profile, users' questionnaires, conception of scenario. Nevertheless [3] and [4] have pointed out lists of heuristics for evaluating 3D IT in VEs. Our EEA system integrates these heuristics.

3. Empirical Evaluation Assistant (EEA)

3.1. Hardware and software context of the EVR@ platform

IBISC Lab. owns a semi-immersive VR/AR platform called EVR@. It permits stereoscopic display, wireless hand/head/fingers tracking and force feedback. Each device is associated to a specific server which is access ed via the C++ *VRPN* library by clients. The interactivity between the user and the VE is done by using *Virtools 4.0* as a front-end. *Virtools* is a good software for prototyping and testing our 3D IT because it offers a fast and graphical way to compute them and link them with hardware devices and VEs by connecting specific building blocks to each other.

3.2. Specification of the EEA

EEA is intended to be used during the development lifecycle of 3D IT by non experts of ergonomics. Typically, it is dedicated to 3D IT for developers. It has no aim to bypass a complete evaluation process made by ergonomics experts.

The main objectives of the EEA system are:

1. To assist experimenters *before the experiment* :
    - o Help for selection of pertinent variables to be traced during the experiment and submitted to statistical analysis after the experiment (correlation detection, hypothesis testing by using ANOVA);
    - o Help for selection of known or personalized protocols to be applied in the experiment;
    - o Help for selection of known or personalized qualitative questionnaires given to the users.

    This assistance is carried out by using a *database* which centralizes the knowledge about 3D IT evaluation: pertinent variables, experimental protocols or questionnaires.
2. To assist experimenters *during the experiment* :
    - o Trace of pre-selected variables during the whole experiment in a log file;
    - o Real time display of pre-selected variables.

    The aim is to permit an *easy debugging* and to detect erroneous behaviors of the users.
3. To produce a feedback about the studied IT *after the experiment* :
    - o Results of statistical analysis made over traced quantitative variables and qualitative variables (questionnaires);
    - o Possibility to replay the experiment off-line;
    - o Integration of the whole experiment results in the *knowledge database*.
4. To permit collaborative work over the *knowledge database*:
    - o Reusing
    - o Share the results of IT experiments with experts outside the Lab;
    - o Annotate the experiments.

3.3. Software architecture of the EEA

In order to achieve the specification of the EEA system, we have built two distinct tools and utilized existing free software. The global architecture and software implementation is given in figure 1.

The first tool is dedicated to *Experimental Protocol Conception*, which we call *EPC* tool. It includes paragraphs *1* and *4* of our objectives. The *EPC* tool permits the access to the *knowledge database*. We fulfill paragraph *4* by choosing a WEB based architecture centered on an *Apache 2* server. The database is implemented with a *MySql* server which is accessed via *SQL* queries from the *EPC* tool written in *PHP* and *AJAX*.

The second tool is dedicated to *Measurements and Debugging*, which we call *MD* tool. It includes paragraph *2* of our objectives. It has been implemented by making specific *Virtools building blocks* that we call *Probes*. The probes may be connected to building blocks which output has to be measured, traced and displayed in real time. Figure 2 shows four probes connected to the tested 3D IT given in section 4. They permit to measure the duration of a user's experiment and how many mistakes he has made. A *Core* component permits to initialize the measurement schema of all pre-selected variables by using a configuration file created by the *EPC* tool before the experiment (curved arrow in figure 1). It also permits to synchronize the data gathered by the different probes, by using dedicated modules (figure 3). The *Synchronization* and *Wait* modules permit to synchronize probes and core. Probes send synchronization messages to these modules. When the synchronization is done the core launches a module (e.g. "*Speed*" or "*Acceleration*") for computing speed or acceleration of specified object on the virtual environment.

The fulfillment of the objectives of paragraph 3 is done by using the log files produces by *MD* tool during the experiment and configured in the *EPC* tool before the experiment. These files are read by *R software* scripts that produce the results (correlation detection, hypothesis validation).

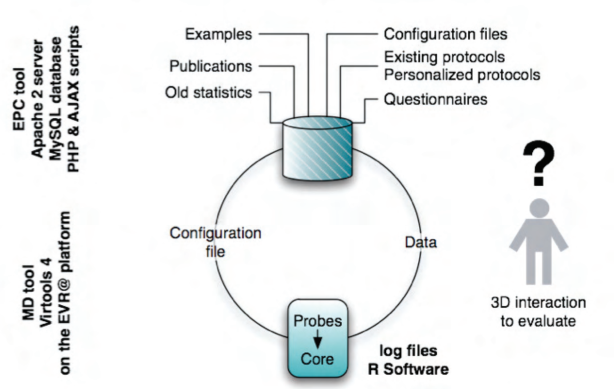


Figure 1. Software architecture and implementation of the EEA system. Statistical analysis is done by the *R* software from the log files generated by the MD tool.

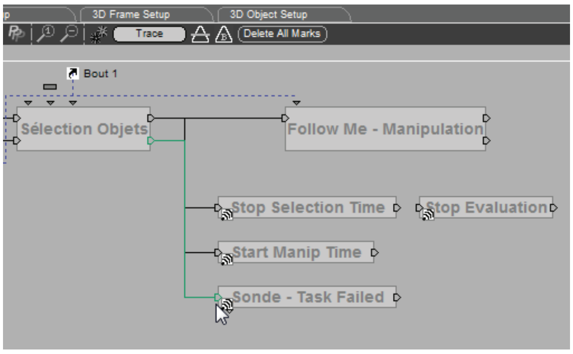


Figure 2. View of four probes connected to a tested 3D IT (*Follow-Me* Manipulation block) in the *Virtools 4* framework.

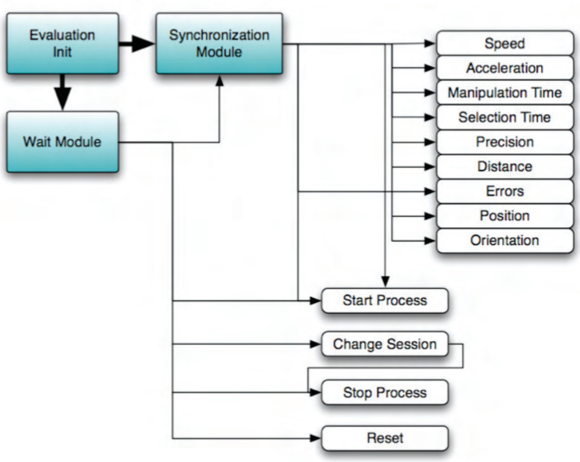


Figure 3. Architecture of the *Core* module of the *MD* tool. The boxes at the right correspond to the pre-selected variables the experimenter wanted to trace in the 3D IT *Follow-Me* testing.

3.4. Experimental protocol design with EPC tool

Three steps are required for creating your own experimental protocol. The first step consists in choosing properly the indicators and variables (main ones are described in [2]). The experimenter is helped during this process. Each parameter (indicators and variables) may be associated with many publications, examples and different help messages. Variables and indicators are gathered in different categories.

The next step is the creation of evaluation scenarios. The experimenter determines the different parameters. The final step concerns the questionnaires. The experimenter has to choose the automatic or semi-automatic mode. The automatic mode permits to create questionnaires directly by the parameters chosen. Semi-Automatic creates questionnaire directly but the experimenter may modify, add or delete questions.

4. Preliminary empirical evaluation with EEA System

We have experienced our EEA system to refine the condition of use of 3D IT we have developed, called *Follow-Me* (see [7] for the *Follow-Me* model). The model had been roughly tested before without the EEA system. However, some results we had were put into questions and some questions remained unanswered.

The particularity of *Follow Me* is its use of *virtual guides* to reduce the number of degrees of freedom in VE when approaching an object to be selected and when manipulating it. Thus, the system *anticipates* what the user may do to lessen his need for concentration. But this system may puzzle the user if its anticipation is wrong. Moreover, we already knew that *Follow-Me* behaves well for selecting far and small objects comparing to classical tested 3D ITs. Our questions were:

- Is there a real benefit of using *Follow-Me* if the object is near from the user in VE?
- Is there a real benefit of using *Follow-Me* for a user who is an expert of VEs?
- How is *Follow-Me* perceived by users (helpful, disturbing and neutral)? Is there any difference in this perception if the user is a novice or an expert?

These questions have implied the creation of specific qualitative questionnaires given to the users after the experiment. It has been done with the help of the *EPC* tool. They also implied the nature of the probes utilized in the experiments which were traced in a log file and statistically analyzed after the experiment (see figure 2 and the right side of figure 3).

Figure 4 shows our experimental setting using the EVR@ semi-immersive platform. We have performed a comparative evaluation of *Follow-Me* and two other classical 3D IT over 15 users. Each user had to select a book on a shelf and put it on another shelf as fast as possible. The device used to interact with the VE was a wireless Flystick which position and orientation were captured by two infrared cameras situated at each side of the wide screen.

Two days of work for one experimenter were necessary to:

- Build and implement the experimental protocol depending on the questions we were asking [the *EPC tool* configures the probes and deliver questionnaires in PDF format];
- Complete the experiment in itself with 15 users (an average of 30 minutes per user was necessary) [*MD* tool produces a dated trace of all probes];
- Analyze the collected data to produce a feedback [dated trace and qualitative data from questionnaires are submitted to a script that uses an ANOVA procedure in the *R* software].

EEA permitted us to know that *Follow Me* is favorably accepted by novices in VE and permits faster selection and manipulation that other 3D IT whereas experts are puzzled by *Follow Me* and prefers classical 3D IT. This feedback will be utilized in the future to refine the use of virtual guides in the *Follow Me* model.

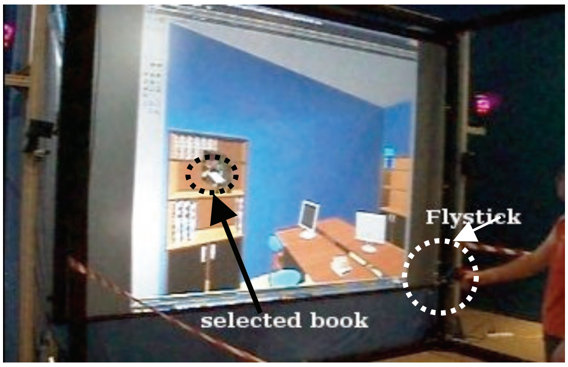


Figure 4. Experimental setting using the semi-immersive EVR@ platform to test the 3D IT *Follow-Me* for a book selection and manipulation task in VE.

5. Conclusion

We have described a tool called *Empirical Evaluation Assistant*. This tool is dedicated to *light* evaluations of Interaction Techniques for VE during their development lifecycle. It may be used by non ergonomic experts. The aims of this tool are:

- A fast design of an experimental protocol by using pre-existing protocols stored in a database;
- Debugging and trace facilities during the experiment;
- Statistical analysis of the inter dependence of pre-selected variables after the experiment.

The core idea is to get fast feedbacks in order to improve the tested 3D IT. In order to accumulate knowledge about our 3D ITs, the whole experiments may be stored into a database which may be accessed worldwide via a WEB interface, whereas the debugging tool is connected to our VR/AR platform and is implemented in Virtools.

We have used our EEA system to test the 3D IT we have developed recently. The feedback we obtained in only two working days permitted us to build an evolution of the *Follow-Me* model.

Future work on the EEA will concentrate on:

- The interface with MATLAB do get statistical analysis online (via MEX codes);
- The collaboration with ergonomic experts to improve our software and share the data collected during the experiments.

Acknowledgments

This work is supported by the *VarSCW* (Virtual and augmented reality Supported Collaborative Work) project. We wish to thank "*Le Conseil Général de l´ Essonne*", *CNRS* and the *IBISC laboratory* for funding this project.

References

[1] Stanney K.M., Mollaghasemi M., Reeves L., Breaux R. and Graeber D.A. "Usability engineering of virtual environments (VEs): identifying multiple criteria that drive effective VE system design", Int. J. Human-Computer Studies 58, pp. 447–481, Elsevier Science, 2003.

[2] Bowman D., Kruijff E., Laviola J. and Poupyrev I. "3D User Interfaces theory and practice", Edition Addison Wesley, USA, 2001.

[3] Gabbard J. L., Hix D. and Swan II J. E., "User-Centered Design and Evaluation of Virtual Environments", IEEE Computer Graphics and Applications, USA, 1999.

[4] Kaur K., Maiden N. and Sutcliffe, A. "Interacting with virtual environments: an evaluation of a model of interaction" Interacting with Computers 11 (4), pp. 403–426, 1999.

[5] Poupyrev I., Weghorst S., Billinghurst M. and Ichikawa T. "A Framework and Testbed for Studying Manipulation Techniques for Immersive VR", ACM Virtual Reality Systems and Technology Conference (ACM VRST97), 1997.

[6] Karaseitanidis I., Amditis A., Patelb H., Sharplesb S., Bekiarisa E., Bullingerc A. and Tromp J. "Evaluation of virtual reality products and applications from individual, organizational and societal perspectives : The ''VIEW'' case study", Int. J. Human-Computer Studies 64, pp. 251–266, Elsevier Science, 2005.

[7] Ouramdane N., Davesne F., Otmane S. and Mallem M. "Follow-Me: A new 3D interaction technique based on virtual guides and granularity of interaction." ACM International Conference on Virtual Reality Continuum and Its Applications (VRCIA 2006), pp. 137--144, Hong Kong, China, 2006.

Haptic Devices Actuation : Functional Overview and Typology

FLORENS Jean-Loup*†
(†)ACROE – INP Grenoble, France*
E-mail:florens@imag.fr

Abstract

This article presents a general analysis on the actuators of haptic devices. This analysis aims first at clarifying the status of this component among the various aspects of haptic devices functions and performances. The functional structure of the actuation system is introduced. This basic structure is then used to establish a general typology that attempts to encompass most of the current and historical haptic technologies. Within the proposed classification, the strength and limitations of the different actuation technologies are discussed.

1 Introduction

The actuator is an essential and critical element of an HD: it constitutes with the mechanical linkages the part of the haptic device that supports the physical constrains/mechanical energy flow that are inherent to the haptic interaction [5].

The haptic device actuator role has been associated to the concept of force feedback and linked to some aspects of the human perception. According to this theoretical guideline the mainframe of haptic actuator performance criteria were based on human tactile-kinaesthetic perception analysis [4]. In addition haptic actuation concept inherited from the classical approach of robotics that situated it in a field of effects and active behaviour. But contrarily to a robot the primary specification of a haptic device is to be a substitute of an object and then to behave mostly passively. The actuator function in this case must be precised since the use of classical actuation concepts may lead to neglect important feature when high precision haptic systems have to be designed.

In the following we firstly examine the functional aspects of the haptic devices actuators and then establish from this model a typology of the actuation systems whose main properties are compared. The objective is to find functional relationships between the existing technical solutions with a quest of genericity for new developments.

2 Functional overview of actuation

In a general functional diagram of an haptic device sensing and actuation blocks are two "ports" elements each provided with one side a mechanical port an on the other side, a signal output or input port

These two components constitute the core of the interface between the user's physical real world and the "informational" domain of the virtual models calculation. The sensor function corresponds to a usual concept of measurement, its produces a signal from its physical interaction with the other parts of the device whose meaning (type of sensed variable) is not a completely intrinsic property of the sensor but is conventional and related to a representation system.

The actuator cannot be considered as the reciprocal of a sensor since it is not possible to set a measurable effect from an input signal without violating a simple causality principle. Then it is more appropriate to consider that the primary function of the actuator input signal is to determine a physical constrain like a temporary structural property. This can be schematically represented by the formula:

$$F(f,v) = T(e)$$

where F represents the instantaneous *(f,v)* relation and T is a temporal operator that represents the causal dependency of this relation on the input.

2.1 The energy flow modulation

In the general case the actuation system is at least constituted of two parts

1) A power supply or "buffer" that produces or absorbs the effective mechanical power of the haptic interaction.

2) A passive modulator that insures the control of the power flow by the input signal. The simpler conceptual component that allows obtaining such a control from a low or null energy input signal is a non-accumulative dipole, i.e. a pure dissipative element.

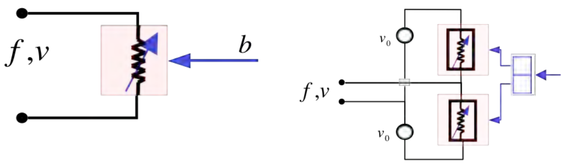
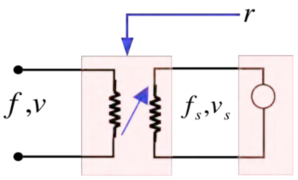


Figure 1. Circuit representation of a controlled brake, and the half bridge push-pull modulator configuration

Then the electric circuit representation of the corresponding actuator structure is a rheostatic circuit whose a symmetric combination allows to cover the four regions of the *(f,v)* plane that correspond to the different combinations of positive and negative power and speed. That leads to the classical half bridge push-pull (Figure 1).

Another modulator type can be obtained from a non dissipative, energy conservative, transformer whose ratio *r(e)* is controlled by an input signal *e*. This quadripole element is interposed between a power supply and the actuator load (Figure 2). In this case the resulting *(f,v)* curve family is the image of the power supply *(f,v)* curve, by a transformation of the controlled impedance ratio *v/f* .



$$f = r.f_s$$
$$v_s = r.v$$

Figure 2. Circuit representation of impedance ratio actuation

In the first case of a dissipative modulator the power supply provides always a positive power. Then the totality of the ingoing power from the physical environment is dissipated by the modulator.

In the case of the impedance ratio modulator the environment power flows is identical to the supply power flow

The two types of modulator present different properties regarding the resulting internal impedance of the actuator. In the first case (a) this impedance is always highly dissipative. In the second case (b) this impedance is an image of the power supply internal impedance and when several modulators share a same supply the different channels are coupled through this supply bus ; this properties is exploited in the cobots for providing a controlled cinematic constrain (§3.4)

2.2 Alternate energy flow modulation

The conceptually simpler actuation system would consists in a complete mechanical realisation of the above schemes. However in the usual context of industrial servo-controlled systems and robotics the technology of such systems could be relatively complex. For this reason, and because of the availability of mature technologies in the domains of energy conversion and power control, alternate energies, mainly electric in lesser extend pneumatic and hydraulic, have been employed in most haptic actuators.

The energy conversion is realized by an additional component at the end of the chain. This component is ideally a non-dissipative and non-accumulative “transducer” Its role is to convey in the most transparent way the physical constrains generated by the electric or fluid modulator into a mechanical constrain. In a electric circuit representation this transducer is a transformer whose ratio defines the relationships that link the electric or fluid variables (voltages/ intensity, pressure/flow) to the mechanical variables (speed/force) .

A particular issue that concerns energy conversion is the commutation: In order to perform the infinite cyclic motions that are natural in the mechanical motion space, the transducer may include a commutation principle. This is one of the basic features of the classical electric DC machines It consists in a mechanical combination of several elementary transducers each operating on a limited portion of the motion space. The commutation system may be independent on the modulator (electric brush motors) or integrated with it (electric brushless motors). Commutation leads to a variation of the actuation characteristic at each commutation step. This drawback named “cogging” is particularly sensitive at low or null speeds.

2.3 Local control loop

In practice the (f,v) curve families of the actuation system are not convenient for the type of control signal that is issued from the calculation. In addition these characteristics may be sensitive to external parameters. To overcome these limitation a local feedback is set up to control the (f,v) /input signal according to the desired type of control (Figure 3).

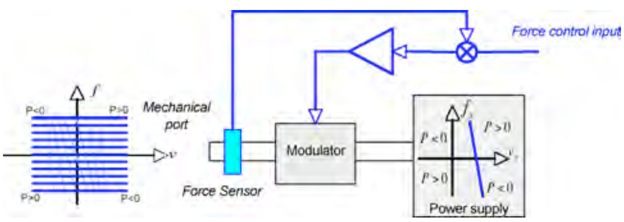


Figure 3 Example of local loop of a force controlled actuator

The use of a control loop leads to a new family of (fv) curves but the availability of these new physical characteristics is limited by the cut-off frequency of the control system, itself depending on the actuator load (the physical environment).

The particular actuators like force actuators or motion actuators are necessarily based on such a local control loop.

3 Typology of actuation

Depending on the type of alternate energy, type of modulator and technology, a wide diversity of actuation systems may exist.

3.1 Electro-magnetic actuators

The electric systems are the most commonly employed because of the suitability of electronics technology in making high performance/ low cost electric modulators. This leads to shift the focus on the electromechanical transducer that is more critical component of the actuation chain. Different physical conversion principles exist. The Laplace force based electro-magnetic devices are the most widespread with DC motors and linear or rotating electro-dynamic transducers. The haptic DC motor are generally commercial devices except some that have specially been developed for high performance haptic/robotic usages[6](§5.3.2). The electro-dynamic systems are specific devices: the ACROE ERGOS shared flux modular multi-axis device [7], the special 2 axis rotating actuator of UW [12] designed for 2 Dofs fingertip haptic.

The essential limitation of the electromagnetic transducers *is quantitative* and related to their maximum specific power and specific force that are determined by the physical properties of the active materials (conductors and magnets)[11]. The force limitation is the most critical for the haptic usage.

Their strengths are their intrinsic linearity, high bandwidth and the obvious advantage of the Laplace force is its ability to circulate through a gap of air thus avoiding material link between the actuated parts. This property that greatly simplifies the structure of the cyclic commuted actuators (the classical DC motors) has also made it possible to design efficient active levitation systems like the CMU Maglev [8].

3.2 Piezzo-electric actuators

Unlike electro-magnetic motors the piezzo-electric transducers provide great forces under low volume. That could be a significant advantage for haptic actuation. On the contrary since they work by deformation, their displacements are restrained. This last limitation is overcome in some miniaturized high torque motors by the use of cyclic operation of several elementary cells with alternate contact /release phases [6]§5.3.6. However these technologies are not still completely convenient for haptics because of the too limited speed and hard non-linear properties of these devices. Finally the haptic uses of piezzo-actuation are mainly limited to the very small working space of tactile stimulators. In this case the actuator consists of sets of independent elementary piezzo-transducers.

3.3 Circulating fluid actuators

The main properties of these types of actuation are described in [6]§5.3.Their common interest is the lightness of the transducer (cylinder) that makes them suitable for portable haptic interfaces[3]. Pneumatic is more convenient for low cost systems. Its major limitations are the poor bandwidth due to the compressibility of the air and the important dry friction of the air cylinders. The hydraulic technology although hard to set-up is potentially highly efficient for haptics

Hydraulic cylinders present high force capability, high bandwidth, high stiffness, and low inertia. Unlike electric actuators the only physical properties that limit the power of a hydraulic converting device are the mechanical limits of the material. As a consequence at equivalent power and force capability the hydraulic device is lighter and smaller than the most efficient electric transducer. Then direct joint actuation is possible thus avoiding gear reduction and commutation.

There is still an open field for adapting the hydraulic technology to haptics and overcoming its technical drawbacks that concerns essentially the availability of high bandwidth servo-valves. Then the hydraulic technology could be an alternate solution to electric actuation for medium and large size systems.

3.4 Mechanical systems

The direct modulation of a mechanical energy flow is possible in two different ways:

1. By controlling the friction coefficient of special clutches or brakes. In this case the mechanical modulator is a push-pull of two such clutches combined with mechanical power supply. The recent advances in smart materials, in particular magneto-rheological fluids is a way to overcome the limits of the traditional electro-magnetic powders in the design of such controllable brakes and clutches [2].

2. By using a continuous variable ratio transmission system as CVT that links the load to a constant speed supply. Such variable ratio transmission systems were employed on an active 6-DoF device called active cobot. This device allows obtaining very high stiffness with low inertia [10].

The limitations of the two types of mechanical modulators are:

1.The use of a constant speed rotating motor is be a source of sensitive vibrations.

2. Their input bandwidth is generally limited by the input electromechanical conversion device.

Their main advantages are:

1. The dynamical limitations inherent to the energy transducer are avoided.

2 Both dissipative and CVT modulators present high dynamical performances in their direct open loop physical properties.

3. The mechanical control of an infinite cyclic motion is natural in the mechanical motion space and does not require any commutation system.

4 Unlike in energy conversion systems the absence conversion losses allows designing efficient passive actuators with the 2 types of modulators

(dissipative & CVT). The power supply is then replaced either by a null motion ground, like in the direct contact MR fluid devices[9], or by a 1dof free motion mechanical buss that allows the passive inter-axis coupling, like in the passive cobots [10].

3.5 Compared properties

Several categorizing properties appear between the different technologies. The first concerns the transfer characteristics related to the control input. The second concerns the physical property the device is able to exhibit out of control or in stationary control mode.

Regarding the haptic requirements the first properties defines the ability of the actuator to follow the closed loop control from the virtual objects calculation. The versatility of the system and the field of variability of the provided behaviour will depend also on this property that refers to the "Z width" concept. The second will define the sizing limits (forces and speeds) of the haptic device and it's out of control behaviour.

Then the mains 3 classes of actuators examined above can be compared according to these criteria: Electric systems present poor intrinsic physical properties (high inertia and limited stiffness) but high control capabilities. On the opposite mechanical system present very interesting physical properties but reduced control bandwidth that limits their usage to specialized application where the desired behaviour fits with their intrinsic physical properties. Hydraulic actuation although rarely used in haptics presents promising features for high performance and large size haptic devices.

4 Conclusion

We have proposed a functional representation of the actuator system with the attempt to clarify its role in the haptic interface functions. From this representation we have proposed a typology of the different technologies encountered in the today haptic devices domain. This typology allows establishing various types of similarities between haptic devices and new criteria for the comparison of the different technologies for a better understanding of the encountered bottlenecks. It appears that a majority of haptic devices are built from the current actuation technologies that had been developed in the various domains of controlled mechanical systems. These devices are generally convenient for large or medium work space where spatial and geometrical constrains are predominant. Several specialized systems are more focused on high performance dynamics or very special usages where specific performances are required (passivity, high bandwidth). These developments have leaded to original actuation solutions whose design is based on a deeper analysis regarding specific haptic issues. Although initially focused on very specific usages these components may constitute alternate solution to the classical systems in the cases where their technology has reached its physical limits. Besides these cases a wide landscape of non explored solution exists either in the field of converting systems or in direct mechanical modulators.

Systematic exploration and comparison of the different actuation technologies could result in a better knowledge of the limits of actuation and provide realistic guideline for future researches in haptic interfaces.

5 References

[1] Ash G. Les capteurs en instrumentation industrielle. Dunod Paris 1991

[2] Bar-Cohen Y., Mavroidis C., Melli-Huber J. and Fisch A. (2002), An Exoskeleton with On-Demand Controlled Resistivity and Operability Using ERF-Based Elements, NASA Tech Briefs, 2002

[3] Bouzit M., Burdea G., Popescu G., Boian R., "The Rutgers Master II – New Design Force-Feedback Glove", IEEE/ASME Trans. on Mechatronics, vol. 7, No.2, June 2002, pp. 256-263

[4] Burdea, G. C. (1996). Force and touch feedback for virtual reality. John Wiley & Sons, Inc.

[5] Cadoz, C. (2004). Enactive interfaces? Enactive Virtual Workshop. http://www.interdisciplines.org/enaction

[6] ENACTIVE (2004). Report on the state of the art on existing sensors/actuators technologies for haptic interfaces. Deliverable D3.1, European Network of Excellence Enactive Interfaces.

[7] Florens, J.-L., Luciani, A., Cadoz, C., and Castagné, N. (2004). ERGOS: A multi-degrees of freedom and versatile force-feedback panoply. In Eurohaptics 2004, pages 356–360, Munich, Germany

[8] Hollis J. R.L., Salcudean S., Allan A.P., "A six degree of freedom magnetically levitated variable compliance fine motion wrist: design, modeling, and control", IEEE Trans. on Robotics and Automation 7, pp. 320-332, June 1991

[9] Michelitsch G.& al. Technical paper, "Haptic Chameleon: A New Concept of Shape-Changing User Interface Controls with Force Feedback," CHI 2004, Vienna, April 24-29, 2004

[10] Moore C. Continuously Variable Transmission for Serial Link Cobot Architectures Thesis for Degree in Master of Science in Mechanical Engineering.

[11] .Trémolet de Lacheisserie E.,& al., Magnetisme , vol. 1 Fondements, Presses Universitaires de Grenoble, ISBN 2.7061.0831.2, pp. 13-28, 1999

[12] Venema S.C., Matthes E., Hannaford E., 'Flat Coil Actuator having Coil Embedded in Linkage,' U.S. Patent N° 6437770, 2002.

Assessment of Global Space Integration by means of Tactile Snapshots of Environment

Eleanor Fontaine* Edwige Pissaloux* Ramiro Velazquez*

(*) Institut des Systmes Intelligents et Robotique (ISIR), France

E-mail: fontaine@robot.jussieu.fr, pissaloux@robot.jussieu.fr, velazquez@robot.jussieu.fr

Abstract

This paper addresses preliminary experimental results on users' abilities for space integration when using a new touch stimulating interface. The aim was to know whether the subjects' were able to integrate different tactile snapshots of an environment into a global representation of it. An experiment involving blindfolded, healthy and naive subjects was lead in order to assess the appropriateness of the interface for space understanding and to integrate it later in a travel aid for the visually impaired. It was conducted using a prototype of a touch stimulating interface based on the Shape Memory Alloy (SMA) technology and designed at the ISIR, Paris 6 University and at CEA/LIST, France. It was found out that subjects have acquired the notion of tactile obstacle, of tactile map and that they were able to represent themselves in an environment.

1. Introduction

Access to information is a key element for man-space interactions like looking, reading, walking, item reaching, navigation. . . [2]. Different access methods are used in function of different supports involved in information representation. Generally, vision is predominant over other senses to get access to information. For the visually impaired, interactions with space are largely deteriorated, leading to difficulties in the fulfilment of daily life tasks. Usually touch stimulating devices are used to display information.

Depending on task different kind of touch stimulating devices are used to supply information. For instance Braille keyboards are mainly used to read and write, while relief maps of public halls are the most popular assistances used to display three dimensional static data (Figure 1). Recent technological developments allow the design of touch stimulating devices providing static and dynamic display (Bach-y-Rita's TDU, Heildelberg's VTD, TeleSensory, Optacon, [4]. . .).

This paper introduces a new touch stimulating surface built with SMA technology at the ISIR through an

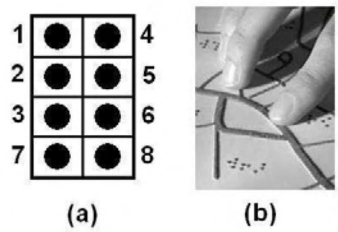


Figure 1: (a) A Braille cell and (b) A relief map.

exprimental evaluation. The final surface's goal is to help the visually impaired during navigation so we first evaluated abilities of voluntary sighted and blindfolded subjects in a virtual navigation task.

The addressed question is whether the subjects' are able, by using the proposed new surface, to understand different tactile representations of an environment and to integrate them to create a global space map.

The paper is organised as follows. Section 2 presents materials and methods including participants, apparatus and protocol. Section 3 summarises the collected results and discuss them. Finally, future research directions are outlined in the Conclusion (Section 4).

2. Materials and methods

2.1. Participants

Five voluntary graduate students from Paris 6 University (4 men and 1 woman) were involved in experiments. They were all healthy, sighted, with no impairments in tactile sensory or cognitive functions, right-handed and with no previous experience in Braille or tactile display usage. Their ages ranged from 24 to 28 years old.

During the experiments, they were sat blindfolded in front of the tactile display, so that no cue from sight could be obtained (Figure 2). Before the session, they were totally naive to all aspects of the test and were given general instructions concerning the task.

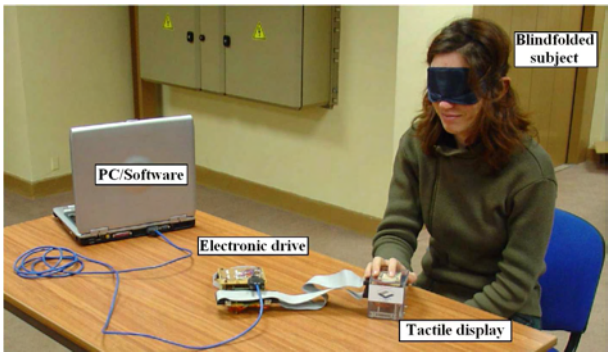


Figure 2: Experimental platform and its three elements: the PC, the electronic drive, the tactile surface.

2.2. Materials and apparatus

A touch stimulating (Braille-like) surface designed and prototyped at Paris 6 University is defined, upon physiological properties of touch sense [1], as a 2D matrix of tactile elements called taxels. Each taxel consists of a metallic pin moving up and down thanks to two antagonist SMA springs (Figure 3); both are fixed using a central mass. SMA springs are grounded at the middle so that the electrical current can flow independently in the upper or lower half [6], [5]. Therefore, raise and retraction of the contact pin is based on the alternatively activation of the halves.

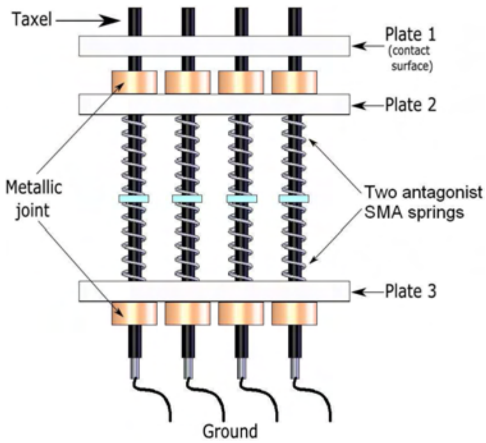


Figure 3: Structure of a taxel.

Figure 4 shows the 8x8 taxel array fully assembled in a plexiglas enclosure. The effective tactile area is 4.32 cm^2 and its laboratory cost is 200 USD. Summarising its key features, this first prototype consists of a 64 element array spaced 2.6 mm apart that vertically actuates SMA based miniature actuators of 1.5 mm diameter to a height range of 1.4 mm with a pull force of 300 mN up to a 1.5 Hz bandwidth. The full display is 200 g weight.

One of the key points is the presence of a notch on the device. This notch is a representation of the subject related to the displayed scene.

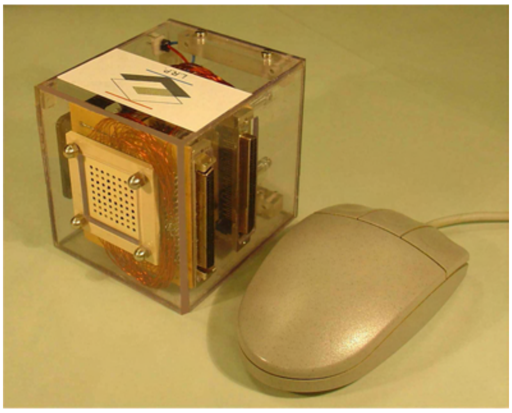


Figure 4: Tactile display in a plexiglass enclosure.

2.3. Protocol

In this experiment we evaluated the subjects' abilities to understand and memorise tactile representations of an environment and to integrate them in order to generate a global space map. The representation of environment was first presented and experimentally validated in [3]. This is a representation of obstacles borders which are the nearest from the subjects.

The first step in the realisation of these experiment was the development of a virtual environment. Figure 5 shows the environment composed by four spaces (called rooms in the rest of the paper). To lead the design of these virtual environment some constraints were used. For instance using empty rooms or using only straight or oblique lines for the composition of each room.

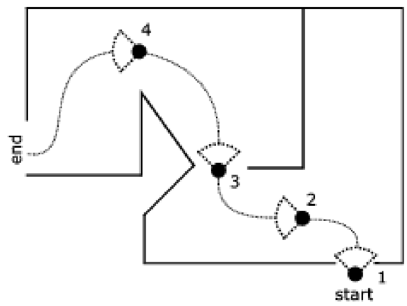


Figure 5: The virtual environment used in. The views are displayed according to the schedule indicated.

2.3.1. Experimental procedure

For each room a view from the entrance is given to the subjects (Figure 5). As the subjects are seated on a chair that means that for each room the tactile representation of walls and free from walls spaces is a static one.

Figure 6 shows these four subject-related views as well as their tactile coding in the 8 x 8 display. Note that for each room there is one entrance and one exit, the exit of one room corresponding to the entrance of the following one. The subjects knew that the exit in a room was going to be the following point of view and the entrance of the next one. The size of exits and entrances are varying according to each room in order to test the capacity of the subjects to recognize an empty space. For simplicity and as a first evaluation, mapping the rooms walls to tactile domain was fixed based on general form and not so much in precise dimensions.

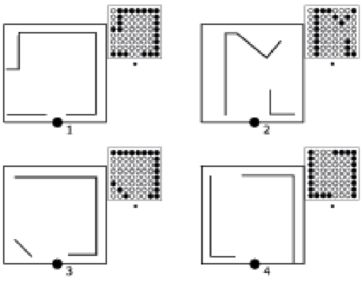


Figure 6: The subjects' views and the corresponding tactile representations. The subject's position in the environment/tactile display is represented by the black dot.

The four rooms were successively displayed on the surface and for each one the subjects were asked to draw it. Finally, the four choices of Figure 7 were presented and they were asked to point out the one they thought to be the navigation environment. Based on the drawings and knowing that the free-from-walls space leads to the "door" to the next room, the four choices were supposed to help the subjects in this task. The three wrong choices were design by modifying difficulties of the correct environment (the triangle and the location of exits).

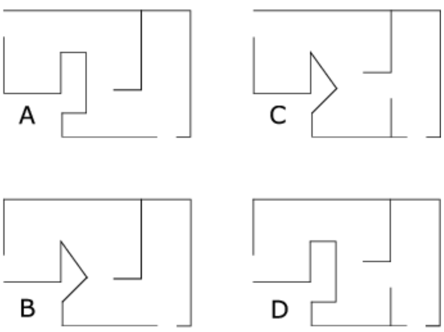


Figure 7: The four environments proposed to the subjects (the correct answer is B).

2.3.2. Instructions to the subjects

A general presentation of the space coding was first made in order to ensure that the subjects clearly understood the representation of walls. Indeed, the representation is a combination of allocentric (walls are represented relatively to the subjects positions) and projective (wall's borders) view. Then, the subjects were invited to touch the notch on the surface that symbolised their position related to the displayed scene. Note that if a displacement would have occurred in the room the notch would have enable the subjects to feel evolutions in borders positions according to them.

Then, the subjects were told that four rooms were successively going to be displayed on the surface and that after having explored the tactile display and considered having fully understood the current tactile map, they were going to report it as a drawing. The subjects explored freely the tactile surface with no time restriction and no specific finger/hand imposed. At the end, each subject obtained four drawings and we asked them to establish the global navigation environment by concatenating them, by using the previous given information concerning entrances and exit in rooms and by seeing the four choices (Figure 7). Each task duration was registered.

3. Results analysis and their discussion

3.1. Quality of drawings

The Figure 8 presents drawings done by the 5 subjects for the four points of view. All subjects' drawings and answers were analyzed in details like for subject-1: in room (1) for the entrance subject- 1 appears to have translated parallel-to-him lines into perpendicular lines. He didn't detect the triangle in room (2) but represented acceptably rooms (3) and (4).

3.2. Environment recognition

3 out of 5 subject were able to recognize successfully their navigation environment by using their drawings. But we conclude that errors in recognizing the environment was totally justified like for subject-1. Thus, considering that 3 subjects recognized the environment we stressed the fact that they were able to represent themselves in the different rooms. We also underline that they got a global representation of rooms' positions related to each other.

3.3. Exploration times

From previous experiments involving the same tactile display and for shape recognition we expected recognition times within one minute but this was rarely observed (Fig. 9). We also observed that the faster subjects just tried to obtain a global idea of the shape explored

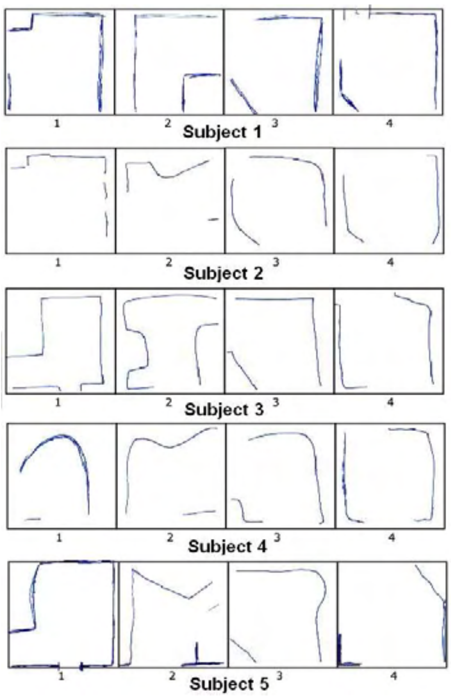


Figure 8: The four views drawn by the five subjects.

while the slower ones sought to understand everything in detail. The global idea was more effective and accurate in all cases. Note also that almost all subjects (excepted subject-2) take less time and become more efficient in exploring the display as the test progresses.

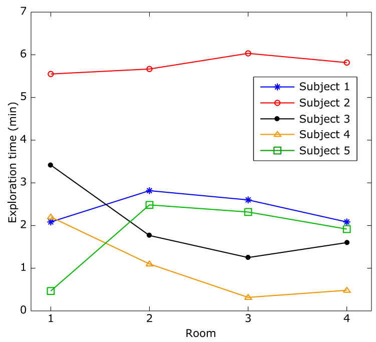


Figure 9: Exploration times per subject for each room during the test.

4. Conclusions

The paper has presented a new touch stimulating SMA based surface and its evaluation via an experiment of tactile snapshots integration.

As expected the proposed space representation seems to be appropriate for space exploration. Indeed, we pointed out that subjects acquired the notion of tactile obstacle (wall) and tactile map during the different rooms explorations. However the spatial representation isn't clear and uniform. The metric data (precise shape and distance) requires more deep psychophysical experiments and probably to test the training effect with the surface.

The obtained results should be compared with the recognition results obtained with piezo-electric Braille keyboards and the same subject population. Other less familiar population with mind-space representations should be include in the study (all the subjects were students in science). But, the main point will be to compare results to those obtained with blind subjects in the same condition in order to know if they are at ease with the space coding.

Future work will involve more formal psychophysical experiments concerned with space organization and space cognitive maps. Moreover, measures for drawing resemblance should be defined in a more formal way. However, these preliminary results obtained from healthy sighted totally naive in tactile display usage are undoubtedly encouraging for testing the system with blind people.

References

[1] M. Bear, B. Connors, and M. A. Paradiso. *Neurosciences, A la découverte du cerveau*. Pradel, 2002.

[2] A. Berthoz. *Le sens du mouvement*. Odile Jacob, Paris, France, 1997.

[3] F. Maingreaud. *Une contribution à la perception de l'espace et son intégration dans une aide pour la navigation des déficient visuels*. PhD thesis, Université Pierre et Marie Curie, Dec. 2005.

[4] R. Velazquez. *Contribution à la conception et à la réalisation d'interface Tactiles Portables pour les déficients visuels*. PhD thesis, Université Pierre et Marie Curie, June 2006.

[5] R. Velazquez, E. Pissaloux, and J. Szewczyk. A low-cost highly-portable tactile display based on shape memory allow micro-actuators. In *Proceedings of the IEEE International Conference On Virtual Environments,Human-Computer Interfaces, And Measurement Systems (VECIMS)*, pages 1344–1349, Giardini Naxos, Italy, July 2005.

[6] R. Velazquez, E. Pissaloux, and J. Szewczyk. Miniature shape memory alloy actuator for tactile binary information display. In *Proceedings of the International Conference on Robotics and Automation (ICRA)*, pages 1344–1349, Barcelona, Spain, July 2005.

Hinge Dimension: An Exploration of Enaction in Architecture

Karmen Franinovic   Alexander Wilson
Zero-Th Association and Parabolik Guerilla
E-mail: karmen@zero-th.org

Abstract

In this paper, we address issues of technologically mediated enactive experience with and within architectural contexts, and the ways in which it may affect our relationships to constructed surroundings. We present the conceptual ideas and design development of the Hinge Dimension project, an interactive architecture which is activated and transformed though the actions of its inhabitants. Changing structural, sonic and luminous arrangements continuously produce different spatial experiences as visitors move though the architecture. We summarize the results that we gathered during its public exhibition, an experiment which was conducted toward understanding the role of enaction in architecture, as well as phenomenological and social issues which emerge in such interactions.

1 Introduction

Several architects have described their work as focused on activating the body and on transferring movement from architecture to it. However, bodily activation though architecture has rarely been done in an instrumental way that allows the visitor to interact with the architecture by touching and manipulating its structure. The architect Lars Spuybroek [4] has used concepts from theories of enaction to design non-tangible media such as sound and to link them directly to the movement of the body. In other cases, the organic and irregular static shapes challenge the body to adapt to unusual environments. However, the physical structure of the space in such cases remains nonetheless immobile and does not allow for its transformation.

Examples from the early 1970s of inflatable architecture challenged structural modifications of the environment from the side of its inhabitants. Groups such as Eventstructures invited people to walk on water, though inflated tubes. Visitors to these installations had to learn to move in new ways in order to navigate the spaces involved. However, decisions as to how the space was changing were not made by its inhabitants. The latter were simply thrown into pre-designed environmental conditions conflicting with those they were used to.

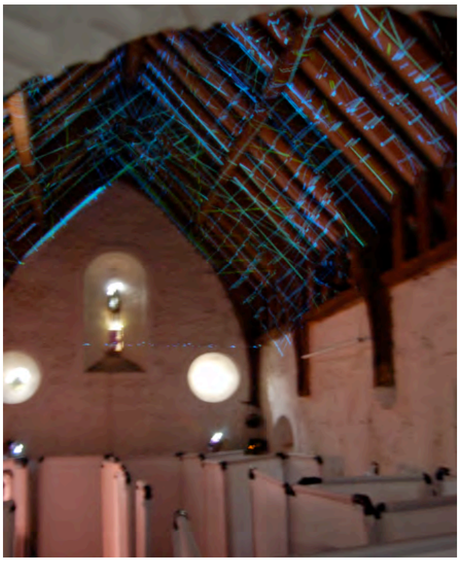


Figure 1: Hinge Dimension: The architecture as arranged close to a linear organization.

In this paper we explore enaction as a process of decision making regarding a spatial structure. The architecture allows visitors to move together with the physical elements composing the environment. The interactive architecture that we describe is not designed as automatically kinetic. Rather it contains a potential for movement and an invitation to its inhabitants to transform its spatial arrangements: structural, sonic and luminous. The actions of its inhabitants are not prescribed and predicable, but simply activated thought the kinetic potential of the architecture.

2 Hinge Dimension

Our movements produce space: as we walk we spatialize the city through the kinesthetic rhythms of our bodies [2]. However, our movements are limited, directed and guided by architecture. Walls direct our

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

movements of the body, values conduct our movements of thought. Can we collectively reconstruct the architectures we inhabit? Can we direct and redirect people moving through their established fixed paths by the means of an interactive system? Can we harness the flow and create monadic zones of intensity in consolidated structures of architecture and thought, such as a church? Hinge Dimension explores these questions and aims at challenging habitual and institutional ways of behaving and inhabiting architecture. The project was commissioned by Enter festival and was exhibited during the 5 days of the festival, which took place in April 2007 in Cambridge (UK).

Originally proposed as an interactive labyrinth for the street environment, a buffer in the urban flow, Hinge Dimension was designed as an architecture in which both soft and hard structures are produced through the action of its inhabitants. Their play with the spatial and bodily relations between themselves and the environment produces transformation formalized through changes of spatial arrangement, light atmosphere and soundscape. Rather than supporting the idea of built architectures as given, immutable structures, this project enables visitors to arrange and re-arrange the tangible and intangible elements of existing space with one gesture. By moving the hinge-screen, at once a doorway and a wall, a connector and a separator, the participants can articulate and redistribute the flows of inertia of fixed architecture. In this way they modify the established uses of the public space and create their own experience of a public area.

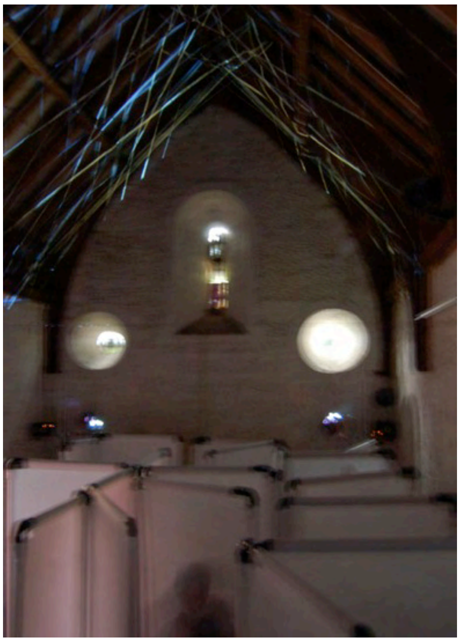


Figure 2: Accumulation of the past activities is visible in the flow of the abstract elements project on the ceiling.

An incessantly changing public space is produced. In it, no motion is automatic, but every voluntary gesture, past and present, is echoed by the responsive environment. The movements of boundaries do not only enclose or open the space, they enable sounds, lights, colors and shadows that intensify space and produce new sensorial atmospheres. Each hinge-screen is a singular articulation of potential movement, but the rotation of this tangible architectural element also modulates light and sound projected back into the created spatial arrangements, accentuating expansion and contractions of a collaboratively built spatial experience. Inhabitants of this architecture co-structure and co-inhabit previously non-existent architectures, the result of their own engagement with visible and invisible flows.

2.1 Technical Description

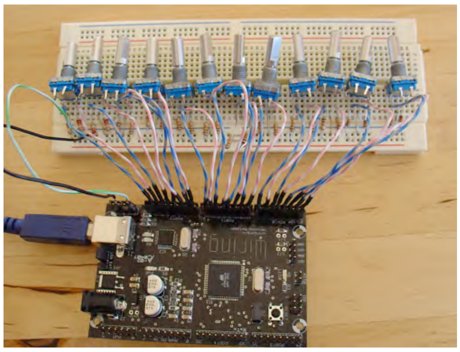


Figure 3: Testing a bank of rotary encoders together with the Wiring microprocessor board.

The physical structure of the Hinge Dimension environment is composed of twenty-eight rotating panels mounted on bases affixed to the large wooden plates on the ground (Figure 1). The latter were interconnected by metal plates, so to form a continuous wooden floor. In this way, a greater stability of the panels was achieved. All electronics and wires were hidden under the wooden floor. The structure of the panels was realized by means of metal and plastic tubes. The panels were dressed in a white fabric that covered most of the structure, leaving only the corners and the bases of the panels exposed.

The metal bases supporting the panels housed continuous rotary encoders, which in turn were connected to inputs of microprocessor boards [1] by the means of shielded cat5 cables. All data reporting the position of the set of rotating panels was transfered to the main computer though two USB connections. Altogether, two microprocessor boards and twenty-eight rotary encoders were employed.

The data received was preprocessed by the microprocessor boards, and transmitted to software running

Figure 4: The single Hinge Dimension screen.

on a host computer. The software was written in the MAX/MSP/Jitter environment, which was used for sensor data processing, sound and video generation. Based on the data that was received, a changing morphology of the space was effected. The overall sonic and visual atmosphere changed according to the number of openings, right angle connections, and the size of the enclosed spaces. Based on resulting inferences as to the organization of the space, the sound sources were spatialized through a four channel sound system, and the ceiling activated by the means of four video projectors. Both audio and visual media reflected the potential for flow of the space in the specific moment. The more immediate response to users action of moving screens was produced though the soundscape, composed of granular sound synthesis processes. The samples used in the latter were recordings of the bodily actions such as swallowing (referencing the link to leprosy noted below) which we previously recorded in the studio in Montreal.

2.2 Adaptability and Site-Specificity

Following the search for an appropriate location, the Festival organizers presented the possibility to install the work in the oldest building in the city of Cambridge. The project's re-situation in a medieval church has strongly affected its development. While interactive and tangible relations between to the architecture and physical structure of the installation did not change, sonic and luminous feedback had to be redesigned for the specific location of the Leper Chapel.

The need for adaptation of the proposed project to a new location permitted us to confront issues related to how interactive systems may deal with the site-specificity. Any situated artwork is designed in relation to its context. Although embedded into a public location, the core structure of an interactive artwork is not necessarily designed as site-specific. It becomes so through the adaptability of the interactive system and through visitors' participation.

The adaptability of Hinge Dimension to different locations is expressed both through its hard modular structure and through the design of the interactive system as an empty box capable of metabolizing different sonic and visual contents. The aim is to provide a context [3] for topological transformation that folds its surroundings into different architectures. Such context must negotiate existing constraints of the location in order to allow for new interactions. These constraints do not only include the physical space, but also its inherent behaviors and associated bodily actions. The ways in which the existing constraints will be challenged depend both on the desire of the audience to play with them and the potential of the system to change.

Soft environmental changes are typical of a number of contemporary interactive artworks, which prefer free movement to tangible engagement with the architecture and artifacts. For example, Camille Utterback has created environments in which ephemerality is perceived through changes of soft media, but leave behind the transformation of the hard matter, making it easy to adapt the interactive artwork to variety of contexts[5].

Figure 5: Leper Chapel in Cambridge

We expanded this idea of adaptability, from the soft media to the tectonic part of the architecture, by designing it as a modular structure that can be easily recomposed and reorganized. Although in our design process we have distinguished between the soft and hard structures of architecture, we see the two as strongly connected through movement and performance. The tectonic, structural changes of the environment, are coupled to soft changes.The simple movement of the vertical boundaries of which the architecture is composed, produces a change in material, social and sensorial architecture.

2.3 Assessment: Visitors' Interactions

As the chapel was displaced from the main festival venue, most visitors were not the festival attendees but people accidentally passing by. We saw this as a benefit, as the variety of audience members was larger. The average number of visitors each day was approximately 80. These numbers were provided by the volunteers from Essex Dance company who welcomed visitors at the entrance of the chapel. The volunteers encouraged visitors to leave comments in a notebook before leaving the installation site. In addition to the comment book, we drew on ethnographic research methods to collect data, namely video recording and approximately 15 informal interviews. Over 5 days, we recorded 12 hours of video documentation and subsequently analyzed it. The following findings are the conclusions from this multifaceted analysis of users' interaction.

Figure 6: Visitor observing the flow of the abstract elements expressing the accumulation of the past activities.

We found that people did not feel invited to touch panels composing the labyrinth. The reasons that they mentioned were: the size of architectural elements, which did not appear as something to be moved; the sacredness of the church space, which made them physically passive; and their expectation that the space would react simply to their presence rather than to their actions upon its architectural composition. In relation to the latter, most visitors expected the space to react to their free movements and did not think of physically engaging with the spatial interfaces.

After the first day, we realized that clarification was needed to indicate that the labyrinth can be touched and rearranged. Still, the interaction was not as large as was expected, and was manifested in smaller movement of the panels, rather than through an overall reorganization of the space. One visitor described her behavior toward the space as a form of "architectural politeness", a kind of fear of large changes that she had to take responsibility for. The direct movement of the large architectural elements, which caused changes of sound and light, appeared to increase awareness of their contribution to the changing space. The obvious visibility of ones physical movement made people shy away from attempting transformation of the environment, engendering the aforementioned "politeness". Children, however, did not evidence such a fear, spending up to an hour and a half playing with the Hinge Dimension.

A number of visitors suggested the size of the labyrinth should be larger. They desired to get lost in the constant transformation of the space caused by themselves and others in the space. Some participants thought the visuals projected on the ceiling of the chapel often revealed too much about the overall spatial organization, allowing them to orient themselves rather than to drift through unknown paths of the labyrinth. However they appreciated the short moments of disorientation.

The reponsivity of the system seemed to be better appreciated by a participants who engaged longer in interaction with the labyrinth. They found that the immediate level of response encouraged them to use the panels as an instrument, and subsequently to understand the overall slower evolution of the sonic and light environment. However, some visitors did not take the time to explore the complex dynamics, and thought that the only response came from a direct mapping of their movement with the architectural panels to the sound and visuals.

The results gathered during the project's opening to the public are valuable toward the further development of the project. The most interesting finding was the fact that people experience problems with the experience of changing the structure of the space and with interacting with large wall-like elements that compose it. At the moment we are exploring ways in which we may engage participants in more vigorous interaction with the physical structure of the architecture. For more information visit http://www.zero-th.org/HingeDimension.html.

3 Acknowledgments

We acknowledge the generous support of the Enter Unknown Territories Festival, Junction, essex dance, Hexagram, Concordia University, and Yon Visell.

References

[1] H. Barragan. Wiring. http://www.wiring.org.co/.

[2] M. de Certeau. *The Practice of Everyday Life*. U California Press., Berkeley, 1984.

[3] N.Gaffney M. Kuzmanovic. *Multiplex Translations / Entangled Aphasia*. http://f0.am, 2001.

[4] L. Spuybroek. Projects' information available at http://www.noxarch.com/.

[5] C. Utterback. Projects' information available at http://www.camilleutterback.com/.

Toward a Phenomenological Pragmatics of Enactive Perception

Tom Froese[*] Adam Spiers[†]
(*) Centre for Computational Neuroscience and Robotics, Uni. of Sussex, UK
(†) Bristol Robotics Laboratory, Uni. of Bristol, UK
E-mail: t.froese@sussex.ac.uk, a.spiers@yahoo.co.uk

Abstract

The enactive approach to perception is generating an extensive amount of interest and debate in the cognitive sciences. One particularly contentious issue has been how best to characterize the perceptual experiences reported by subjects who have mastered the skillful use of a perceptual supplementation (PS) device. This paper argues that this issue cannot be resolved with the use of third-person methodologies alone, but that it requires the development of a phenomenological pragmatics. In particular, it is necessary that the experimenters become skillful in the use of PS devices themselves. The 'Enactive Torch' is proposed as an experimental platform which is cheap, non-intrusive and easy to replicate, so as to enable researchers to corroborate reported experiences with their own phenomenology more easily.

1. Introduction

The enactive approach to perception, which holds that perception consists in perceptually guided action, was first proposed by Varela and colleagues in the early 1990s as an important addition to the embodied-embedded and dynamicist movement in the cognitive sciences (e.g. [16]). Since its inception it has inspired a range of related approaches such as the sensorimotor approach [10], Noë's enactive approach [9], and the dynamic sensorimotor hypothesis [6]. While there are certainly important differences between them, we will refer to these approaches as "enactive" to the extent that their overall concern "is not to determine how some perceiver-independent world is to be recovered; it is, rather, to determine the common principles or lawful linkages between sensory and motor systems that explain how action can be perceptually guided in a perceiver-dependent world" ([16], p. 173).

One important source of evidence supporting this enactive approach to perception is the data generated by experiments with subjects who have mastered the skillful use of a "sensory substitution" device [1]. Such a human-machine interface, which for reasons given by Lenay and colleagues [8] we prefer to call a *perceptual supplementation* (PS) device, is an innovative tool for investigating some of the fundamental mechanisms of perception. A PS device functions by making use of neural plasticity [2], namely the ability of the central nervous system to incorporate the device's feedback into the subject's sensorimotor exploration of the environment [8]. Such a PS device can provide an extremely rich experimental platform, one whose full potential still has to be explored.

The work with tactile-vision substitution systems (TVSS) originated in the early 1960s (e.g. [1]) and continues to be of great practical and theoretical importance in the cognitive sciences [2]. A blind (or blindfolded) user of a TVSS can successfully engage in tasks which normally require visual perception such as navigating novel environments, locating objects in space, reading text, as well as distinguishing between different persons based on their external appearance [1]. Accordingly, research with TVSS has important practical implications for the development of better prosthetic visual systems. Furthermore, such research provides significant empirical support for the enactive approach since it highlights the importance of action in perception (e.g. [9,6,2,10]). It has been observed that the capacity for action is needed for exercising the appropriate sensorimotor skills which are constitutive of our perceptual abilities [1]. It is also necessary for the acquisition of these skills since "only through *self*-movement can one *test* and so *learn* the relevant patterns of sensorimotor dependence" ([9], p. 13).

Interestingly, a subject's mastery of skillfully using a TVSS device is often reported to result in the constitution of a specific spatial mode of perceptual awareness. However, there is little agreement about the best way to characterize the phenomenology of this perceptual modality [2]. Is it an extension to an existing modality such as touch [3] or vision [9,10,6]? Or is it, perhaps, the constitution of an entirely new perceptual modality [8]? The arguments are generally based on two sources of evidence: (i) the experimenter's descriptions of the *abilities* of the subjects who master the skillful use of the apparatus (e.g. to explore their environment while avoiding obstacles), and (ii) the

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

experimenter's descriptions of the *verbal reports* which the subjects provide of their experience. Since both of these sources provide third-person data, the debate about the character of a subject's experience is consistent with one popular third-person approach to the scientific study of conscious experience, namely Dennett's *heterophenomenology* (e.g. [4]).

In this context the aim of this paper is twofold: (i) to argue that the current debate in the literature provides a good example of why such a purely third-person approach is not sufficient to achieve clarity on these kinds of first-person issues, and (ii) that in order to promote a *phenomenological pragmatics* (e.g. [14,15]) of enactive perception, namely a research program which goes beyond a purely third-person approach in a principled manner, PS devices need to be standardized and made readily available to the general research community. Accordingly, the current third-person approach to the phenomenology of TVSS use is critically reviewed (section 2), the development of a first-person approach is motivated theoretically, and a design for a simple PS device that is cheap, non-intrusive and easily replicated is outlined (section 3).

2. A case study in heterophenomenology

One of the most contentious issues surrounding the enactive approach to perception is the question whether the experience constituted by the skillful use of a prosthetic visual system such as TVSS is actually *visual* in character (e.g. [10,9,12,3,2]). How should this be decided? There is general agreement that blind (or blindfolded) subjects must have some sort of *spatial awareness*, since they are able to navigate their environment (at least to some extent). However, since this is a matter of determining the characteristics of first-person phenomenology, Prinz [12] makes the important point that a description of the subject's *abilities* is not sufficient evidence to settle this matter.

What about the published descriptions of the *verbal reports* provided by the subjects? One of the problems here is that they allow contradicting interpretations. Moreover, the reader of the literature generally has no access to the actual reports, but must base her interpretation on the descriptions provided by the experimenter. As an example consider the following account in an early seminal paper by Bach-y-Rita and colleagues [1]: "our subjects spontaneously report the external localization of stimuli, in that sensory information seems to come from in front of the camera, rather than from the vibrotactors on their back". O'Regan and Noë [10] interpret such descriptions as evidence that the experience of the subjects is similar to vision, and Noë ([9], p. 27) claims that "it is reasonable to admit that the resulting experiences are, if not fully visual, then vision-like to some extent". Block [3], on the other hand, argues that "there is doubt as to whether the phenomenology of TVSS is exclusively visual" and that "perhaps TVSS is a case of spatial perception via tactile sensation". Similarly, Prinz [12] seriously doubts whether subjects experience anything visual: "My best guess is that prosthetic vision devices simple allow subjects to make automatic inferences about where objects are located in space as a result of tactile information". Again, it appears that the published reports are not sufficient to settle this issue.

So far the problem of determining the qualitative nature of the experience associated with TVSS use has been limited to the interpretation of descriptions of behavioral or physiological data and verbal reports, and evidently without much success. Can this issue be resolved by these sources of evidence? This would seem to be the hope for anyone advocating a purely third-person science of human consciousness such as Dennett's heterophenomenology [4].

However, even if it is assumed that such an approach can be made internally consistent, an assumption which will be questioned in the next section, there remains a practical problem. It is all too easy to be content with simple textual interpretation when the principled analysis of first-person experience is not assigned any explicit role [14]. This seems to be the case in the current debate about the phenomenology of TVSS use, where a scientific investigation has slowly been turned into an open-ended debate about mere interpretations of interpretations. Indeed, there is no evidence that any of those involved in the debate have had experience with a TVSS themselves. In this respect Prinz [12] has to be commended for at least making this issue explicit. He openly discloses that he has not used Bach-y-Rita's apparatus himself, and therefore admits being forced to venture a "best guess" on what its use could be like. It is doubtful that a consistently third-person approach could ever go beyond postulating such educated guesses while at the same time remaining true to its foundational principles.

Is a purely heterophenomenological approach even possible in the first place? It has been pointed out by Gallagher [5] that any such third-person study of consciousness is not entirely free of phenomenological elements. We similarly claim that such studies are implicitly based on the first-person experience of the investigator herself in two fundamental ways: (i) first-person experience *in general* is presupposed by any (theoretical) activity, and (ii) first-person experience of the *particular* experiential phenomenon being investigated is presupposed by any meaningful interpretation of the third-person data.

Presupposition (i) is basically a reformulation of the irreducibility of consciousness as the necessary background which frames all of our activities ([16], p. 9-12). In this context it means that rejecting the existence of first-person experience in the researcher outright makes any attempt to understand first-person experience intrinsically self-refuting. It eliminates that which affords the possibility of the attempt itself – the authentic nature of awareness and purpose. Moreover, even if the investigator were such a hypothetical disembodied and purely rational intellect she could only (if it all) wonder why the objects under study make certain sound patterns and move in a particular way. It is because of *our own* first-person experience of being conscious subjects that it is possible for us to even conceive of investigating how other subjects undergo a certain experience and pick out the relevant third-person data. For a more detailed argument along these lines see Jonas ([7], pp. 127-134).

This kind of self-refutation would also be a problem for Dennett's heterophenomenology if it consistently applied the *intentional stance* to all conscious subjects. The investigator and the other 'subjects' would then be turned into objects that behave *as if* they were subjects but which would be considered as nothing more than "theorists' fictions" [4]. Of course, this problem can be avoided by exempting the researcher from applying the intentional stance to herself. Still, this concession is a first step toward a full blown phenomenological approach. The question is then not whether first-person experience plays a role *per se*, but rather *what kind* of role it plays. This brings us to presupposition (ii).

The debate about the experience associated with skillful TVSS use has generally not been informed by personal experience with the device by any of those involved in the dispute. Nevertheless, the arguments are clearly based on some aspects of their first-person experience of perception. How else is it possible to argue about whether the TVSS experience might be vision-like and or a form of spatial perception via tactile sensation? A strict heterophenomenology would be like demanding that only congenitally blind researchers can investigate the phenomenology of vision. It is clear that they would be immensely aided in their task if they themselves had experience of what having vision is like. Similarly, Dennett [4] concedes that "it has always been good practice for scientists" to try out their own experimental apparatus. However, as soon as this has been explicitly recognized, the natural question is: why not also provide them with a more principled manner of corroborating verbal reports with their own experience? Hence, this second concession opens the door for acknowledging the importance of a full blown *phenomenological pragmatics*.

Of course, this kind of first-person approach to the scientific study of consciousness also acknowledges the importance of third-person data. Indeed, the goal of this change in perspective is the establishment of a relationship of *generative mutual constraints* between phenomenology and cognitive science (e.g. [14,5]). In this manner the research community is explicitly encouraged to bring all available sources of evidence to bear on the problem of consciousness.

3. Toward a phenomenological pragmatics

The existence of first-person experience in the researcher is a legitimate source of additional evidence which not only needs to be explicitly acknowledged but also practically cultivated [15]. It carries with it the responsibility of making a concerted effort to engage in disciplined training to describe experience accurately [14]. Moreover, undergoing appropriate guidance will also enable researchers to become more skilled at obtaining the relevant phenomenological descriptions of untrained subjects as well [11]. This kind of 'second-person' approach is in stark contrast to heterophenomenology, which explicitly rejects communication in favor of detached interpretation and thereby diminishes its means for clarification of meaning [5]. That such training not only enhances the ability to describe one's experience accurately but also changes the experience itself is not a problem but actually an inherent necessity. Indeed, only by attending more closely to one's experience, and thereby undergoing a different kind of experience, can one describe it more accurately [11]. The original experience is not lost – it is just brought into focus.

The biggest hurdle toward the establishment of such a phenomenological pragmatics is that it requires some radical re-learning from the research community [14]. This is an important topic but beyond the scope of this paper (but see [15]). What is significant in this context is that researchers must be able to train in the use of a PS device. This is also important for the insertion of the subject into a world of shared meaning [8] on the basis of a relationship of mutual trust [11]. Unfortunately, such an endeavor is often impractical when dealing with expensive and complex commercial products. A better starting point would be a cheap PS device which is effortlessly replicated, simple to handle, non-intrusive, and for which the design can easily be made available to the public. Such a PS device is the Enactive Torch (ET).

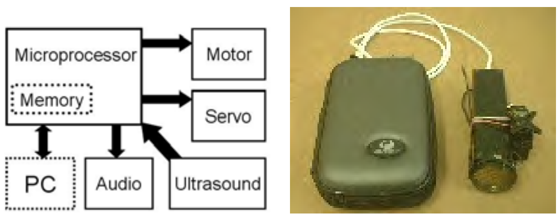


Figure 1. A schematic of the enactive torch (left) and an image of a recent experimental prototype (right)

The ET provides the subject with one continuous channel of vibro-tactile feedback, where the strength of stimulation depends on the distance to the object which is currently pointed at. As shown in Fig. 1, the ET also contains a servo-motor and audio hardware to which the output may be routed if the researcher desires. The design is based on an earlier PS device, the Haptic Torch, which has been successfully used by blind subjects to navigate simple environments [13]. Its practical advantage over a traditional cane is that it avoids the intrusiveness of direct contact and thus encourages exploratory movements, especially in social contexts. This is important considering the role of action in the deployment of such a new mode of perception [8].

Almost immediately a subject can use the ET to detect obstacles, and after a little training (ca. 10mins) has no problem of locating relatively ‘transparent’ objects, such as a lamppost, in space. Moreover, already after around 1hr of practice certain salient features of the environment, such as corners and open doors, take on a distinctive perceptual pattern which could be described as ‘touching objects out there’. Appropriate sensorimotor coupling can thus give rise to perceptual resolutions that significantly exceed those of single-point stimulation. Moreover, we hypothesize that severing the causal relationship between action and tactile sensation when using the ET, for example by playing back a recording of a previous trial, will prevent the constitution of an experience of spatial exteriority. We expect that under such conditions subjects will be unable to make sense of their surroundings, even though the ‘information’ that is potentially provided by the tactile sensations will be the same as before. Initial trials with short training times (ca. 30mins) have indicated this to be the case. These results are in accordance with the claim of the enactive approach that perceiving is not about the recovery of a perceiver-independent world, but rather consists in perceptually guided action. The ET thus provides the means to establish a mutually informing relationship with a study in the phenomenology of the senses such as that done by Jonas ([7], pp. 135-152).

Interestingly, these early results are comparable to those obtained with a similar PS device by Lenay and colleagues [8]. This suggests that the constituted perceptual modality is relatively independent from the particular kind of hardware implementation and more dependent on the type of sensorimotor coupling which it makes available. Such intersubjective validation provides further support for the enactive approach.

In general, the establishment of a phenomenological pragmatics of enactive perception would be immensely aided by the creation of a common web repository where the research community can post transcriptions of verbal reports as well as technical diagrams of the PS devices used. This is particularly important as the experiential data is related to the technology used [8].

4. Conclusions

The establishment of a proper research program in enactive perception requires a phenomenological pragmatics in which researchers are themselves skillful in the use of PS devices. Initial trials with the Enactive Torch indicate that it could provide an appropriate starting point for this important endeavor.

References

[1] Bach-y-Rita, P., Collins, C.C., Saunders, F.A., White, B. & Scadden, L. (1969), “Vision Substitution by Tactile Image Projection”, *Nature*, **221**:963-964

[2] Bach-y-Rita, P. & Kercel, S.W. (2003), “Sensory substitution and the human-machine interface”, *TICS*, **7**: 541-546

[3] Block, N. (2003), “Tactile sensation via spatial perception”, *TICS*, **7**: 285-286

[4] Dennett, D.C. (2003), “Who’s on first? Hetero-phenomenology explained”, *JCS*, **10**: 19-30

[5] Gallagher, S. (1997), “Mutual enlightenment: Recent phenomenology in cognitive science”, *JCS*, **4**: 195-214

[6] Hurley, S. & Noë, A. (2003), “Neural Plasticity and Consciousness”, *Biology and Philosophy*, **18**: 131-168

[7] Jonas, H. (1966), *The Phenomenon of Life*, Evanston, Illinois: Northwestern University Press, 2001

[8] Lenay, C., Gapenne, O., Hanneton, S., Marque, C. & Genouëlle, C. (2003), “Sensory Substitution: Limits and Perspectives”, in Y. Hatwell *et al.* (eds.), *Touching for Knowing*, Amsterdam: John Benjamins, pp. 275-292

[9] Noë, A. (2004), *Action in Perception*, The MIT Press

[10] O’Regan, J.K. & Noë, A. (2001), “A sensorimotor account of vision and visual consciousness”, *BBS*, **24**: 939-1031

[11] Petitmengin, C. (2006), “Describing one’s subjective experience in the second person: An interview method for the science of consciousness”, *PCS*¸ **5**: 229-269

[12] Prinz, J. (2006), “Putting the Brakes on Enactive Perception”, *Psyche*, **12**: 1-19

[13] Spiers, A. & Harwin, W. (2004), “The Haptic Torch: A Torch for the Blind”, *Int. Conf. on Disability, Virtual Reality and Assistive Technology*, Oxford, June 2004

[14] Varela, F.J. (1997), “The Naturalization of Phenomenology as the Transcendence of Nature: Searching for generative mutual constraints”, *Alter: Revue de Phénoménologie*, **5**: 355-381

[15] Varela, F.J. & Shear, J. (1999), ‘First-person Methodologies: What, Why, How?’, *JCS*, **6**: 1-14

[16] Varela, F.J, Thompson, E. & Rosch, E. (1991), *The Embodied Mind*, Cambridge, MA: The MIT Press

Postural suprapostural task and specific experience

Gautier, G.* Thouvarecq, R.*
(*)CETAPS EA3832, University of Rouen, France
E-mail: geoffroy.gautier@etu.univ-rouen.fr, regis.thouvarecq@univ-rouen.fr

Abstract

The aim of this study was to investigate the effect of a specific experience, such as gymnastics, associated with the effect of a suprapostural task on the postural control. Participants had to maintain the erect posture while looking at a target located in front of their head or their feet. A postural perturbation was applied to test the kinematic adaptations. The peak-to-peak angular amplitudes of the ankles, knees and hips relative to this perturbation were analysed. The results showed an effect of expertise and of the target position on the postural suprapostural task. It confirms the influence of the suprapostural task on the postural control. Moreover, a specific experience leads to a modification of the postural facilitation modes

1. Introduction

Human's possibilities for action (affordances) emerge from the interactions between the environment, the task and the organism's properties [3]. Erect posture can be one of these affordances and the same environment can have different affordances for different organisms. Posture itself also affords many other behaviours. Posture is not an end in itself but rather it supports the accomplishment of suprapostural tasks [5]. More specifically, the postural control seems to be more determined by the suprapostural task influence than by the environmental conditions [4]. However, the interaction between the organism's properties and the suprapostural task remains unclear.

Moreover, a stable postural state corresponds to a minimisation of uncontrolled movements in the perception-action coupling when destabilisation occurs [3]. Uncontrolled movements indicate the loss of stability which hinders the successful performance of suprapostural tasks.

Since expertise in gymnastics is widely acknowledged to modify the organism's properties [1] and to influence the postural control [2], we expected, by comparing gymnasts' and non gymnasts' behaviours, to determine the interaction between the organism and the task. A postural perturbation was imposed to understand how the posture is modulated to preserve the success of the suprapostural task after this perturbation. It allows to investigate the postural synergies according to both the prior experience and the suprapostural task.

2. Method

Two groups of 5 experts in gymnastics and 5 non gymnasts were asked to stand barefoot while looking at a target. The target was a black square (2×2cm) placed either on a white wall, at a distance of 2m from participant' head or on the floor, 40cm from their feet.

A casting weight representing 7% of their weight was fixed in the back of the participants (Figure 1). The height of the weight support was adapted to each participant's height to ensure the same anterior-posterior constraint parallel to the ground. The postural perturbation was obtained by removing the weight without informing the participant.

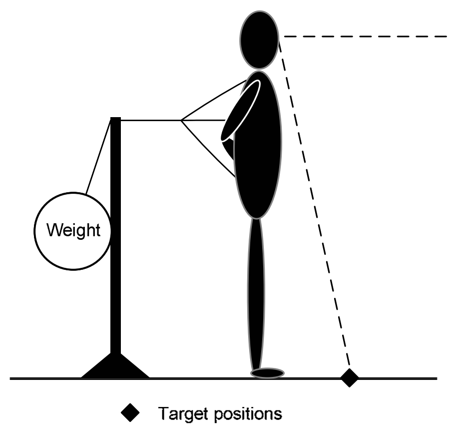


Figure 1. Representation of the experimental device.

One trial per condition was realised and all trials were filmed in profile. Five reference markers were fixed on the right side of the participants, on the small toe, ankle, knee, hip and shoulder. The peak to peak amplitudes of the ankle, the knee and the hip angles were analysed to identify the postural adaptations following the perturbation.

A 2 group × 2 target position ANOVA with repeated measures on the second factor was computed. Finally, pairwise comparisons were performed between mean values using Tukey's post-hoc tests. The threshold for significant differences was set at $p = .05$.

3. Results

Means and standard errors of the angular amplitudes are presented in the Table 1. Data were signed to identify a closing (-) or an opening (+) of the angle.

For the hips, significant effects of group ($F_{(1,8)} = 28.57$, $p < 0.001$), of the target position ($F_{(1,8)} = 22.11$, $p = 0.001$), and of interaction ($F_{(1,8)} = 7.57$, $p = 0.02$) were observed. Post-hoc revealed that, compared to the gymnasts, angular amplitudes of the non gymnasts were significantly greater when the target was on the wall ($p = 0.003$) (Figure 2). More, gymnasts exhibited significant greater angular amplitudes of the hips when the target was on the floor than when it was on the wall ($p = 0.003$).

For the knees, a significant effect of group ($F_{(1,8)} = 86.66$, $p < 0.001$) was observed. However, no significant target position ($F_{(1,8)} = 0.001$, $p = 0.99$) and interaction effect ($F_{(1,8)} = 1.61$, $p = 0.24$) was noted. Post-hoc tests showed that gymnasts exhibits significant greater angular amplitudes of the knees than the non gymnasts whatever the condition ($p < 0.003$).

Table 1. Means (standard errors) of the angular amplitudes (°) for the 3 joints according to the group and the suprapostural task

	Hips	Knees	Ankles
Target on the wall			
Gymnasts	-16.80	37.17	-31.80
	(7.61)	(8.81)	(7.44)
Non gymnasts	-43.77	3.29	29.07
	(7.65)	(2.71)	(5.28)
Target on the floor			
Gymnasts	-43.55	32.86	-24.35
	(8.76)	(10.93)	(4.56)
Non gymnasts	-50.77	7.68	27.72
	(6.16)	(3.67)	(7.84)

For the ankles, no significant effect of group ($F_{(1,8)} = 1.60$, $p = 0.25$), target position ($F_{(1,8)} = 1.13$, $p = 0.32$) and interaction ($F_{(1,8)} = 2.34$, $p = 0.16$) appeared.

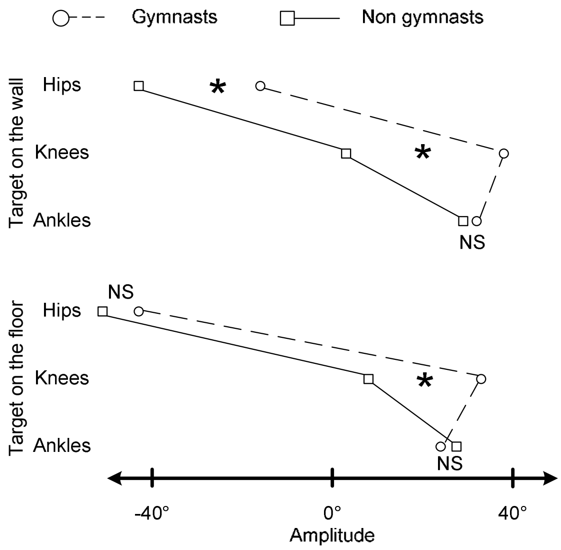


Figure 2. Mean angular amplitudes for the 2 groups in the 2 conditions. * corresponds to a significant intergroup difference and conversely for NS.

4. Discussion

First, our results revealed that the postural control depends on the prior experience. Gymnasts exhibit different modes of postural adaptation than non gymnasts (Figure 3).

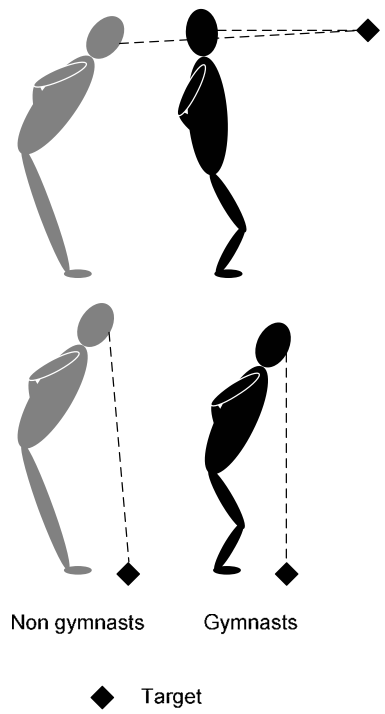


Figure 3. Illustrations of the postural modulations adopted by the 2 groups to facilitate the visual task.

The non gymnasts adopt a hip-ankle synergy to control the postural suprapostural task as those previously observed in studies on the postural control [2]. More particularly, the knees are solicited by the gymnasts which decrease the height of the body mass relative to the support surface. It seems to facilitate the control of the perturbations consequences. Indeed, lowering the centre of mass allows to decrease the anterior-posterior axis of rotation and hence the moment of inertia.

Then, these results are in line with previous findings: the postural control is modulated to facilitate the suprapostural task [4,5]. Indeed, gymnasts, for example, used the hips to keep the gaze on the target when it was located on the floor. Finally, although the expertise modifies the patterns of facilitation, the suprapostural task super-determines the behaviours.

References

[1] Bringoux, L., Marin, L., Nougier, V., Barraud, P.A. and Raphel, C. Effects of gymnastics expertise on the perception of body orientation in the pitch dimension, Journal of vestibular research, 10, pp. 251-258, 2000.

[2] Marin, L., Bardy, B.G. and Bootsma, R.J. Level of gymnastic skill as an intrinsic constraint on postural coordination, Journal of Sports Science, 17, pp. 615-626, 1999.

[3] Riccio, G.E. and Stoffregen, T.A. Affordances as constraints on the control of stance, Human Movement Science, 7, pp. 265-300, 1988.

[4] Stoffregen, T.A., Smart, L.J., Bardy, B.G. and Pagulayan, R.J. Postural stabilisation of looking, Journal Of Experimental Psychology. Human Perception And Performance, 25, pp. 1641-1658, 1999.

[5] Stoffregen, T.A., Pagulayan, R.J., Bardy, B.G. and Hettinger, L.J. Modulating postural control to facilitate visual performance, Human Movement Science, 19, pp. 203-220, 2000.

Annotation: A Support for Co-interpretation

Guyet Thomas
CNRS/TIMC & LIG, Grenoble, France
E-mail: thomas.guyet@imag.fr

Abstract

Clinician are interested in interpreting biomedical time series data to understand patient records. This paper proposes a new co-interpretation approach based on a "structural coupling" between two autonomous agents (a clinician and a machine) to support clinicians in their interpretation task.

From an Enactive point of view, the interpretation can not be autonomously performed with direct knowledge exchanges. Then, annotation seems to be a core concept to enable mutual understanding and to preserve the required autonomy of each interpreter.

The CATS system has been developed to be coupled with a clinician. It autonomously annotates time series data and evolves through its annotation experience.

1. Interpretation

In the large range of interpretation assumptions of Iser [8], there is interpretation where there is a "gap between what is to be interpreted and the register into which it is translated" (p. 83). The interpret had to bridge the gap between the data and the meaning that he can construct from it. There is no predefined meaning in the data. The meaning emerges for the interpreter in his own world of meaning constructed from his past experience.

Therefore the interpreter is seen as an Enactive system. Varela et *al.* [12] define an Enactive system as a system that brings forth a world while it is built by it. We can raise a link between this definition and the case of a world of meaning. The interpreter elucidates meaning from the data and at the same time refines its own perceptions and knowledge.

We separate perception and knowledge refinements to highlight that the meaning is caught in a vice between the uncontextual meaning creation and the meaning appropriation circular processes.

Meaning construction is a forms/meaning issue. The meaning has to be abstracted from the data based on characteristic features (forms) identified by the interpreter. Then the meaning is dependent on the perception capabilities of the interpreter, but perceptions are also modified by the extracted meaning by focusing the interpret attention. Consequently, the interpretation progressively makes consistent forms and meaning in a circular movement.

Meaning appropriation contextualized the meaning in construction in the interpreter knowledge acquired from past experiences. It involves the interpreter in another circular movement in which his knowledge is enriched by new meaning and at the same time it constraints the meaning elucidation.

To resume, the interpretation is the result of the interpreter course along the two indefinite nested circles in which the interpreter is involved to construct the meaning and to appropriate it.

2. Collaborative interpretation

The co-interpretation is a common interpretation performed by two agents. The expected aim of a co-interpretation is to benefit from two different points of view on the data and consequently to improve the interpretation performed by individual interpreter (*i.e.* make converge the interpretation, reduce the gap between the data and the meaning or speed up the interpretation).

In our application, the collaborative interpretation is performed by two cooperating agents, a man and a machine. A human expert is recognized as able to take accurate decisions in rather complex situations, by integrating a wide range of contextual information and keeping a global outlook over the data at hand. Conversely, the machine is able to process large amounts of data under complex numerical constraints.

The collaboration "occurs when two or more agents work together in a common environment to more effectively reach the maximal union of their goal" [2]. In [4], the authors define the main characteristics expected for man-machine collaboration system based on the study of man-man cooperation. Following these authors, the system must:

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

- operate within an acceptable framework of coordination,
- be able to recognize and accept the collaborator's goals when declared,
- be able to interactively work toward super ordinate goals in solving complex tasks,
- offer alternative solutions to the problem addressed,
- operate to support the formation of new attitudes (adaptation).

The mixed-initiative approach [7] appears to date as the more elaborated tentative to achieve these characteristics. In a mixed-initiative system both the system and the user have balanced contributions for problem solving.

In the same vein, we propose an approach based on a "structural coupling" between man and machine to reach the collaboration. The man and machine "may become reciprocally structurally coupled through their reciprocal selection of plastic structural changes during their history of interactions. In such a case, the structurally plastic changes of state of one system become perturbations for the other, and vice versa, in a manner that establishes an interlocked, mutually selecting, mutually triggering domain of state trajectories" [9]. We point out three differences between a mixed-initiative approach and the proposed structural coupling approach:

- the role of our system is not to recognize the human needs in order to assist her/him, but to contribute in a balanced way to the solution,
- in our system man and machine share a common goal,
- the adaptation during the problem solving process is central in the structural coupling approach.

In this context, interpretation is not considered as a context-free attribution of meaning, but rather as grounded in each agent's experience. There is no prevalence of one agent on the other, rather, there is a possibility of learning and discovery for both agents. This kind of partnership is rather meant to allow a co-construction of meaning, in which the interpretation of facts is not defined beforehand by one of the partners, but co-constructed in the course of their interaction.

Such a structural coupling cooperation raises the question of how the two agents can share knowledge without reducing the structural coupling substance. In fact, the knowledge can not be directly exchanged, otherwise agents do not work anymore in their own world of meaning and then the co-interpretation process is no more enriched from various points of view and the structural coupling between agents collapses. Annotation appears a core concept to cope with this difficult issue.

3. Annotations

An annotation is defined by Bringay et *al.* [3] as "a particular semiotic content linked to a target. The target can be a collection of documents, a document, a segment of document or another annotation. Each annotation has a content, materialised by an inscription. It is a trace of the mental representation elaborated by the annotator about the target. The content of the annotation can be interpreted by another reader. The anchor links the annotation to the target".

An annotation is linked to a target and then it is a part of the virtual environment (constituted by the data and the annotations) shared by the two agents.

In a way similar to the talking heads of Steels [10], which interact to build a shared lexicon, based on their independent perception, analysis and manipulation of geometrical figures, we propose to consider both the clinician and the machine as agents who share a common virtual environment, and mutually interact using annotations manipulations to progressively refine their interpretation.

Annotation receives a growing interest in the co-design field and has been shown to support the dynamics of co-operation [1]. Annotations may be seen as tangible marks that can be managed by the agents, *i.e.* they enable the co-construction of objects. They may also be seen as tangible signs that make sense, *i.e.* they are the materialization of contextual knowledge that may be shared among the agents. According to this principle, each agent is in turn given the possibility to observe and interpret annotations provided by its agent, and/or to propose annotations judged as appropriate according to a given interpretation focus.

Annotation supports the co-interpretation based on a structural coupling in two dimensions. Firstly, the interpretation via annotation is "enacted" by agents in the way that they recreate the meaning in their own world of meaning following the interpretation process presented in section 1. Secondly, interactions through annotations preserve the interpretation autonomy of each agent and thus allow establishing the expected "structural coupling"

4. CATS: Collaborative Annotation of Time Series

The CATS system has been developed to support the interpretation by clinicians of time series data (Records of patients in intensive care units). The system autonomously performs the interpretation and interacts with the clinician via annotations.

4.1. Time series data interpretation

The interpretation of time series data consists in extracting significant events and scenarios (events combinations). Based on these significant entities we use three annotation types :

- segments that delineates interesting parts of time series,
- symbols that associates segments to events,
- relations between symbols to annotate time series with scenarios.

For mutual understanding, the semiotic content of our annotations is reduced to the minimum. In particular, there is no textual content that a machine can not appropriate.

4.2. Global view of patient records co-interpretation process

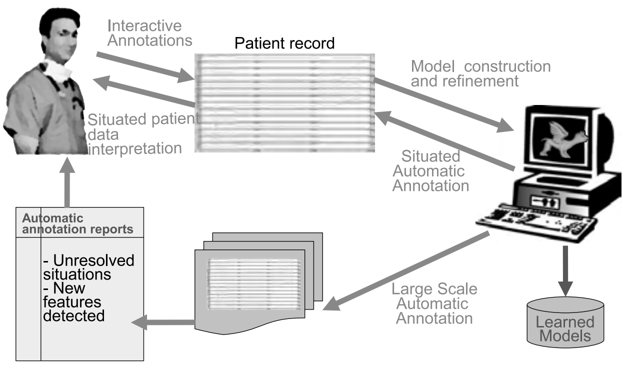


Figure 1: Global view of the interpretation process.

The figure 1 illustrates the global view of the co-interpretation process. The clinician annotates a patient record to interpret the clinical case. The interpretation is situated by contextual information that describes the patient pathology, etc. At the same time, the system builds its models from clinicians annotations and performs automatic annotations on the same patient record (co-interpretation). In exchange, machine models and annotations allow the clinician to refine its own knowledge.

Moreover, the machine performs in parallel large scale annotation of all patient records. Then, machine difficulties or new features identified are reported to the clinician for future situated patient records co-interpretations.

4.3. Autonomous system description

The CATS System autonomously performs the interpretation of the time series across four successive abstraction levels : time series, segmented time series, symbolic time series and situations (record annotated with scenarios). In parallel, the system builds models of events and scenarios. The models constitute the knowledge acquired by the machine along the interpretation and are used to perform the annotations. Consequently, the machine annotation evolves according to its “experience”. Some feedback processes make consistent the new acquired meaning and the knowledge and thus simulate the appropriation of the meaning by the machine. More details about the system can be found in [6].

Due to feedbacks processes and dynamic insertions of annotations by the human, the knowledge construction is a highly dynamic process. Therefore, CATS is a multi-agents system. Agents dedicated to each processing step have been constructed:

- Segmentation agents perform the segmentation. Each agent embed an event model and browses time series to find patterns that match its model and then annotates the time series with the corresponding segments.
- Classification agent collects segments and classifies them. Classes of segments are identified with events and in a second step, the classification agent translates segments in symbols. The symbolic name of a symbol is associated with an event.
- Learning agents collect the resulting symbolic time series. Each agent aims to explain occurrences of an event. The explanation is a scenario learned using a APriori-like algorithm [5]. In a second step, the agent performs scenario recognition to annotate time series with situations.

4.4. Processing example

The figure 2 illustrates a co-interpretation of time-series via segment and symbols annotations. (a) Two annotations (horizontal bar) are inserted on the signal by the clinician to indicate *asynchrony* periods. (b): Annotation completion by the system: dark-gray boxes indicate retrieved *asynchrony* periods. Medium-gray boxes indicate retrieved *non-asynchrony* periods (shorter than *asynchrony*). (c) The clinician has annotated the unrecognized segments. (d) A new model for new periods has been discovered (single light-gray box), the “asynchrony” model has been refined to retrieve all *asynchrony* periods.

5. Evaluation strategy

Our evaluation framework is based on three levels: feasibility, performance and usability. Because we propose a new approach for collaboration, we should evaluate its feasibility. Feasibility means that our implemented system: 1) provides an effective structural cou-

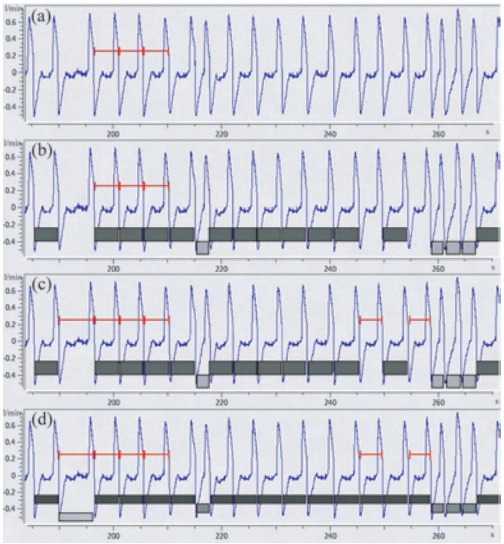


Figure 2: Co-interpretation via annotation examples (See details in section 4.4).

pling collaboration preserving the autonomy and adaptability properties of the system. Its capabilities for models construction, models emergence and automatic annotation should be highlighted; and 2) shows interesting characteristics compare to other collaborative approach.

For time series data exploration, the system performances can be assessed considering three aspects: (1) the capability to efficiently explore and annotate a large amount of time series data, (2) the quality of the built computerized models of events and scenarios, (3) the capability of new events and scenarios discovery. With simulated data sets, objective measures of quality can be proposed. Then, we can compare performances between a fully, partial or absent collaboration with various systems.

The usability evaluation includes human-machine interaction criteria and cognitive science criteria (work load evaluation, result confidence, etc.). It requires the use of interviews and questionnaires to gather experiment feedbacks. In term of acceptability by the users, the adequacy of the collaboration in the clinical environment and gains compared to the standard practice should be evaluated via clinical trials.

6. Conclusions and perspectives

Structural coupling and Enaction theory appears to be a new interesting approach for the collaborative interpretation of time series data. In particular, they lead us to cope with the mutual understanding issue with exchanges of annotations. Annotation seems to us an support of co-interpretation in the sense that it allow preserving autonomy of each interpreter and facilitate the appropriation of the co-constructed interpretation of time series.

The perspective of this work is the evaluation of our CATS system constructed on autonomous annotation and evolution principles.

References

[1] Aubry, S., Thouvenin, I., Lenne, D., and Okawa, S., “Knowledge integration for annotating in virtual environments”, International Journal on Product Development, 4(6):533-546, 2007.

[2] Black, J., Robertson, S., and Zachary, W., Cognition, Computing, and Cooperation, Intellect Edition, 1990.

[3] Bringay, S., Barry, C., and Charlet, J., “Annotations for the collaboration of the health professionals”, Proceedings of AMIA Annual Symposium, pp.91-95, 2006.

[4] Clarke, A., and Smyth, M., “A co-operative computer based on the principles of human cooperation”, Journal on Man-Machine studies, 38(1):3-22, 1993.

[5] Dousson, C., and Duong, T.V., “Discovering Chronicles with Numerical Time Constraints from Alarm Logs for Monitoring Dynamic Systems”, Proceedings of the International Joint Conference of Artificial Intelligence, Morgan Kaufmann, pp.620-626, 1999.

[6] Guyet, T., Garbay, C., and Dojat, M., “A Human-Machine Cooperative Approach for Time Series Data Interpretation”, Proceedings of the 11th Conference on Artificial Intelligence In Medicine, Springer-Verlag, pp.3-12, 2007.

[7] Horvitz, E., “Principles of Mixed-Initiative User Interfaces”, Proceedings of CHI’99, ACM SIGCHI Conference on Human Factors in Computing Systems, pp. 159-166, 1999.

[8] Iser, W., The Range of Interpretation, Columbia University Press, 2000.

[9] Maturana, HR., “Biology of language: The epistemology of reality”, In Psychology and Biology of Language and Thought: Essays in Honor of Eric Lenneberg, Miller, G.A., and Lenneberg, E. editors, Academic Press, pp.27-63, 1978.

[10] Steels, L., and Kaplan, F., “Collective Learning and Semiotic Dynamics”. Advances in Artificial Life. Proceedings of the Fifth European Conference on Artificial Life, pp. 679688, 1999.

[11] Subrahmanian, E., and Monarch, I., “Annotations and Collaboration: One in Service of the Other”, Proceedings of the International Workshop on Annotation for Collaboration - Methods, Tools and Practices, pp.73-82, 2005.

[12] Varela, F., Thomson, E., and Rosch, E., L’inscription corporelle de l’esprit, Seuil, 1993.

Evaluating Two Novel Tactile Feedback Devices

Thomas Hulin* Philipp Kremer* Robert Scheibe† Simon Schaetzle* Carsten Preusche*

(*) Institute of Robotics and Mechatronics, DLR (German Aerospace Center), Germany
(†) Advanced Realtime Tracking GmbH, Germany
E-mail: thomas.hulin@dlr.de, robert.scheibe@ar-tracking.de

Abstract

Adding haptic feedback to virtual reality simulations and telepresence applications increases immersion to human users. As for some applications the use of force-feedback devices is impracticable or not sufficient, tactile devices can be utilized as alternative or as supplement. This work investigates two new tactile devices, which use different technologies for generating tactile feedback. These devices are evaluated by experts for two applications, for a virtual reality simulation, and for teleoperating a hand-arm system.

1 Introduction

Haptic feedback can be split into kinesthetic and tactile feedback. A tactile display is a human-system interface, that generates tactile stimuli for human skin. There are tactile interfaces for several applications, like teleoperation, laboratory analysis, sensory substitution, surface generation, braille systems, and games, see [1].

Recently, two new tactile devices were introduced, i.e. (A) the A.R.T. tactile finger feedback device [6] and (B) the DLR vibro-tactile feedback device for the human arm [5], illustrated in Fig. 1. The former displays tactile information to finger tips, the latter generates vibrational feedback to the forearm. Both were designed to increase immersion in virtual simulations.

This paper gives a short technical description of these two devices. It introduces two possible applications in which these devices can be used, namely haptic VR-simulations and teleoperating a robotic hand-arm system. These systems have been evaluated by a group of experts.

2 System Description

The two tactile devices that are investigated in this publication are based on different techniques for generating tactile feedback. This section gives a technical description of the devices' functionality (for a short overview see table 1).

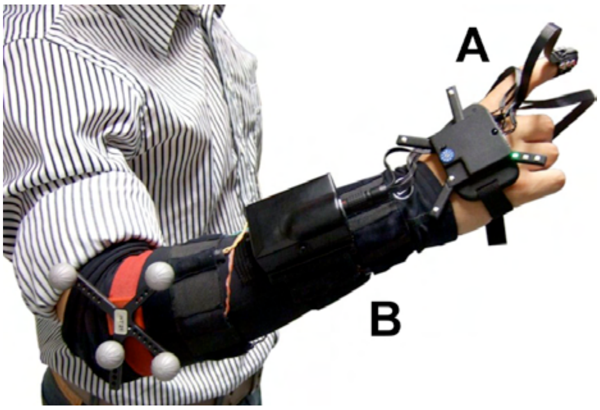


Figure 1: Two tactile devices on a human hand and arm: (A) The finger tracking device with tactile feedback at the finger tips, which is a collaborative work of A.R.T. and Volkswagen, and (B) the DLR vibro-tactile feedback device.

2.1 A.R.T. Tactile Finger Feedback Device

The A.R.T. tactile finger feedback device (FFD) shown in Fig. 2 is extending the A.R.T. finger tracking device with tactile feedback at the finger tips [6]. It is tracked by an A.R.T. optical tracking system. To this end, the device is equipped with active infrared markers (LEDs) for three fingers and the back of the hand. The LEDs for the fingers are mounted on thimbles, which are opened on the lower side. The maximum update rate for tracking the hand target is 60 Hz. As the three finger LEDs are activated time-sequentially, the update rate for each is 20 Hz.

To provide tactile feedback, three shape memory alloy wires are wrapped around each thimble, Fig. 3. These wires have a diameter of 80 μm, and a length of 50 mm. The state of each wire can vary between

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

Figure 2: The A.R.T. tactile finger feedback device (FFD).

a contracted and a relaxed configuration, depending on its temperature. They are activated using a pulse width modulation (PWM) signal. The contracted state is achieved for high temperature $T > 65°C$ using a high duty-cycle. With lower frequency PWM signals also vibrations can be induced to the wires. Therefore, two different kinds of tactile stimuli are possible: short and intense contractions, and vibrations. The communication to the device is wireless using a zigbee radio transmitter.

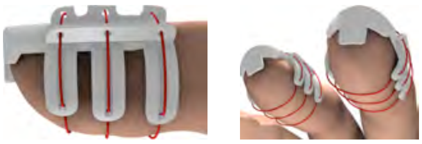


Figure 3: Schematic view of FFD thimbles.

2.2 DLR Vibro-Tactile Feedback Device

The second tactile device investigated is the DLR vibro-tactile feedback device (VTD) [5]. The purpose of this device is to generate vibro-tactile stimuli on a human forearm. It is equipped with 12 groups of vibration motors, each containing two cylindrical motors that are rotating in opposite direction. The motor groups are arranged in two rings around the forearm, see Fig. 4. In a previous evaluation [5] this arrangement has turned out to be a good compromise between resolution and perceivability of stimuli. Each motor group can be controlled separately. The vibration speed of the motors increases with applied voltage and reaches its safe maximum at a voltage of $7\,V$ a frequency of more than $250\,Hz$. The time constant of the motors employed is around $50\,ms$.

Table 1: Features of the tactile feedback devices

FFD	VTD
Tactile feedback at finger tips	Tactile feedback to human arm
Shape memory alloys generate tactile feedback	24 vibration motors in 12 groups generate tactile feedback
Wireless data communication	USB Connection
Short delay in tactile feedback ($50\,ms$)	Motor time constant of about $50\,ms$
Built-in active markers for optical tracking	Built with standard components
Adaptable for individual finger sizes	Variable intensity possible

The device is provided with USB connector for fast and easy connection to a computer. Communication and motor control is an order of magnitude faster than the motor dynamics.

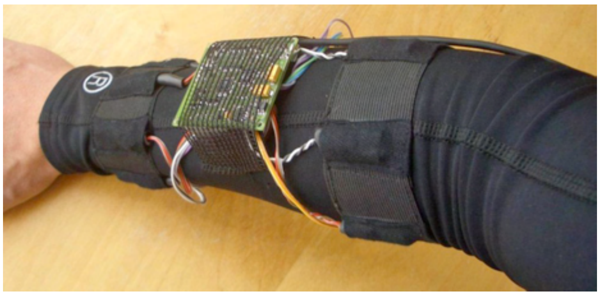


Figure 4: The DLR vibro-tactile feedback device (VTD).

3 Applications

Typical scenarios for human-system interfaces are virtual reality (VR) simulations and telerobotics. For both scenarios an application is presented in this section using one or both of the investigated new tactile devices. Both devices are used to navigate in virtual worlds and feed back collision information. In a second application the FFD is used to command a robotic hand-arm system and feed back information on encountered forces.

3.1 Tactile Devices for Haptic VR-Simulations

A first application, for which the two tactile devices are examined, is tactile feedback from VR. The used tactile

Figure 5: A human operator is evaluating the two tactile devices by exploring a virtual model of the German Micro-Satellite BIRD. The pose of the human arm and fingers are tracked optically by A.R.T. infrared cameras and visualized in the virtual world (yellow cylinder and hand at the top right corner).

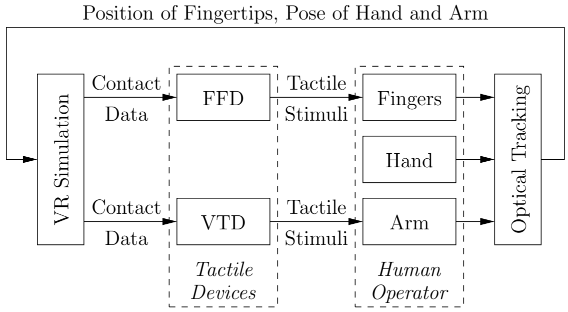


Figure 6: Functional diagram of the tactile VR-setup: Collisions between the virtual representation of the human arm and the virtual scene is sent to the two tactile devices. They activate their actuators to generate tactile stimuli for the human operator. The reaction of the operator is measured by an optical tracking system and the pose of virtual representation of the human arm is updated.

devices enable a human operator to explore a virtual scenario while perceiving collisions. This kind of simulation is used for several purposes, psychological studies, training of mechanics, assembly verification, and gaming.

The employed virtual scenario consists of a virtual model of the German satellite BIRD, Fig. 5. The human hand is tracked optically using A.R.T. optical tracking and the active markers of the FFD. As the VTD has no tracking equipment included, an optical marker is put on the elbow of the human arm. Thus, the human is able to move his forearm, hand and fingers inside the virtual world.

While moving, both tactile devices display concurrently collision information from a virtual environment to the human operator, Fig. 6. Thus, the human operator is able to perceive contacts with the finger tips as well as collisions of the virtual arm; he/she can explore the virtual scene. As the VTD consists of several vibration motor groups, the operator is also perceiving the location of collisions on his arm.

3.2 Finger Feedback for Teleoperated Robotic Hand-Arm System

In a second application the FFD is used in combination with a robotic hand-arm system, in order to telemanipulate this robotic system, Fig. 7. Telerobotic systems are necessary for operating in distant, unreachable, or hazardous environments, for example telerobotic systems in space for on-orbit servicing, minimal invasive surgery, or maintenance of nuclear plants.

The employed hand-arm system is composed of a DLR robot hand II [2] mounted on a DLR light weight robot III (LWR) [3]. The robotic system is coupled to a human operator through positioning the robotic hand and arm according to the measurements of the tracked human hand, Fig. 8. Hence, the pose of the human hand base is used to command the LWR, and movements of the human fingers are scaled and added to an initial pose of the robotic fingers. Position scaling is necessary due to the different sizes of the human and robotic hand. The unmapped fourth finger of the robotic hand is kept in a secure position. All position signals from the FFD are filtered to take care of noise and different sampling times (the robotic system is running at 1 kHz).

Tactile feedback information is generated from the error between commanded and measured positions of the robotic hand's fingertips. Therefore, the operator is able to perceive contacts of the robotic hand.

4 Experimental Results

For both applications an experimental setup was built and tested by eight experts from the field of robotics and/or teleoperation. These experts were interviewed afterwards to obtain their qualified opinion.

For the VR application the human operators were able to explore a virtual satellite with their hands and arms, while perceiving contacts through the two tactile devices. In the experiments the VR simulation including tactile feedback was compared to pure VR visualization. All subjects stated, that immersion was increased for the simulation where tactile feedback was present. Especially the concurrent use of both devices improved this

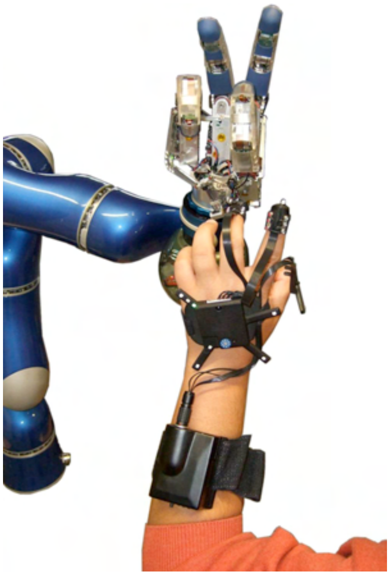


Figure 7: The DLR Robotic Systems mimics the movement of the human operator. The operator perceives contacts of the robotic hand.

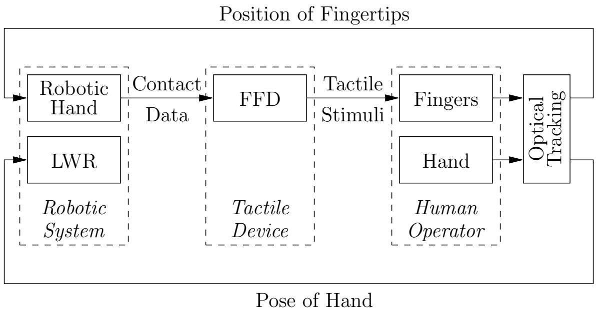


Figure 8: Functional diagram of the tactile teleoperation setup: Collision forces of the robotic fingers are sent to the tactile device. It activates its actuators to generate tactile stimuli for the human operator. The reaction of the operator is measured by an optical tracking system, to update the configuration of the robotic system.

impression for the implemented VR application. Furthermore, all subjects agreed, that for both devices contact impulses are perceived more prominently than persisting contacts or sliding on surfaces.

In the second application, the robotic hand-arm system could be moved in free space and the hand was used to grasp several objects, stiff and soft ones. The contacts were displayed by the FFD, but the actual contact force could only be estimated by visual inspection of the hand or the object. Grasping, placing, and manipulating an object was possible using this human-system interface. In order to do sophisticated manipulations, which the hand is capable of, a more diverse feedback is necessary.

5 Conclusions

Two different tactile devices have been evaluated for two applications, a virtual reality simulation, and teleoperating a robotic hand-arm system. Each device has several advantages. The DLR vibro-tactile device is capable of displaying tactile feedback on 12 different locations to a human arm and can be commanded with a high sampling rate. The A.R.T. tactile finger feedback device is wireless and has already integrated active markers which allow for accurate optical tracking.

Both devices have demonstrated to increase immersion in VR. The FFD can be used as a telerobotic human-system interface. As both devices give tactile feedback to different parts of human skin, it is sensible to use them concurrently for tactile applications.

For future work it should be evaluated if the two tactile device are also suited for manipulation tasks inside VR scenarios, e.g. virtual assembly verification [4]. Furthermore, it seems promising to compare different strategies of actuator control, for instance emulating continuing contacts by pulsing stimuli or using intermediate levels of exertion for the shape memory alloys instead of only turning on and off.

References

[1] Benali-Khoudja, M., Hafez, M., Alexandre, J.-M., Kheddar, A. ”Tactile interfaces. A state-of-the-art survey”, 35th International Symposium on Robotics (ISR’04), pp. 721–726, March 2004.

[2] Butterfaß, J., Grebenstein, M., Liu, H., and Hirzinger, G. ”DLR-Hand II: Next Generation of a Dextrous Robot Hand”, ICRA Int. Conference on Robotics and Automation, Seoul, Korea, May 2001.

[3] Hirzinger, G., Albu-Schäffer, A., Hähnle, M., Schaefer, I., and Sporer, N.: ”On a New Generation of Torque Controlled Light-Weight Robots”, ICRA Int. Conference on Robotics and Automation, Seoul, Korea, May 2001.

[4] Hulin, T., Preusche, C., and Hirzinger, G. ”Haptic Rendering for Virtual Assembly Verification”, World-Haptics Conference 2005, Pisa, Italy, March 2005.

[5] Schätzle, S., Hulin, T., Preusche, C., and Hirzinger, G. ”Evaluation of Vibro-Tactile Feedback to the Human Arm”, EuroHaptics, Paris, France, July 2006.

[6] Scheibe, R., Moehring, M., and Froehlich, B. ”Tactile Feedback at the Finger Tips for Improved Direct Interaction in Immersive Environments”, 3DUI Symposium, Charlotte NC, USA, March 2007.

Video-based Physical Model Animation

Christian Jacquemin*
(*) LIMSI-CNRS & Univ. Paris 11, France
E-mail: FirstName.Name@limsi.fr

Abstract

The design of Allegra, a new type of instrument for animated graphics is presented. It relies on the cohesive coupling of video image analysis and mass-spring system animation. The system allows for direct manipulation through the modification of a 3D scene or through a live video of the user. GPU-based physics computation enables large scale real-time mass spring system animation that makes the user feel like manipulating a fluid or an elastic matter.

1. Introduction

Thanks to the increasing power of Graphic Processing Units (GPU), two domains of physical modelling can be implemented in real-time: physical models for sound synthesis [7] and computer-based painting [1]. These two domains have two important common features. On the rendering side, they both use physical models to imitate natural phenomena: the vibration of a musical instrument or the diffusion of a viscous flow. On the interaction side, they both can be controlled by gestures and allow for direct manipulation.

Our work proposes to combine these two lines of research by adapting the mass-spring systems (MSS) used as physical models for music synthesis to real-time visual effects for video. It could be considered as "Yet another special effect for VJ'ing". But we see *Allegra* as a new type of interactive instrument for video-based graphical and sonic effects that can be controlled by a pre-recorded video or by a performer. When compared with previous work on particle-based artwork such as *Will.0.w1sp* [11] or Physical Modelling Library for Pure Data (PMPD) [9], our approach enables more complex data structures and dynamics. Such a scale factor in size and complexity results in richer graphical effects that makes the users or spectators feel like manipulating extensible organic matters. For this reason, our system has been named after *Allegra Geller*, the main character in David Cronenberg's film *eXistenZ* and designer of organic mixed reality games.

2. Gesture-Based and Physically Modelled Interactive Art

Real-time interactive music systems pay much attention to the role of gesture in musical expression [2]. The coupling between perception and action is expected to enhance affective involvement and artistic expressiveness. Gesture interaction generally involves two components: a capture that can analyze and decode human gesture and a real-time simulator that can provide the artist with a powerful and controllable musical output [5]. Physical models are natural candidates for the design of a gesture-controlled music synthesizer. One of the earliest implementation of interactive physical models was the CORDIS-ANIMA model in the early 80s.

On the side of interactive graphics, physical modelling is used in the domain of Non-Photorealistic Animation and Rendering (NPAR) to simulate the behavior of natural drawing and painting media. These models, such as *Wet and Sticky*, implement the physical and behavioral characteristics of paint rather than just its color properties [4]. The physical models behind computer painting are based on equations for viscous flow. (Fluid dynamics models also have applications in interactive graphics for the simulation of diffusive media such as smoke [10].) Conversely physical models for music are generally based on MSS and greater attention is paid to gesture control than for real-time graphics.

The work presented in this article concerns the design of *Allegra*, a generic and integrated environment, that realizes the mapping between gestures and a physical model in a cohesive and engaging manner. Its purpose is to bridge the gap between interactive graphics and music through the design of real-time multimedia and multi-modal environments. *Allegra* is a video-based physical animation tool that combines interactive graphics with the MSS used in physical models for sound synthesis. It uses simple video image analysis techniques to associate gestures with real-time graphics. Animation is predefined through prerecorded videos, or direct manipulation is enabled through live video or interactable 3D scenes. A previous work on merging graphic and audio renderings in a cohesive physical world was focused on audio-visual feedback [8] and relied on MSS as resonators in an audio-graphic feedback loop. Even though this system was smaller in scale than *Allegra*, its belongs to the same line of research that tries to find tighter connections between graphic and musical worlds.

3. Video-based Control of Physical models

The purpose of *Allegra* and its implementation is to offer an intuitive and efficient framework for video animation of physical models that accepts multiple video sources and various physical models. Physical animation and image processing are both performed in the GPU in order to allow for interactive rendering even with large scale physical systems and complex image analysis. In the current implementation, the physical model used for video-based animation is made of mass-spring chains that are bound by one or both extremities to an anchor in the video image (typically a bright pixel). No collision detection is performed and the dynamics of the physical model is linear. Frame rate is 70 FPS for 3000 chains of an average length 12 with a 3GHz processor and a Nvidia GeForce 7800.

MSSs are composed of particles with a mass and connected by springs. The behavior of the masses is defined by their interaction with the other masses: the force applied to a mass is the sum of the elasticity forces over all the springs connected to this mass (computed from Hooke's law).

$$\vec{F} = -\sum k\vec{l} \tag{1}$$

An additional viscuous damping coefficient adds a force proportional and opposed to the mass velocity. It results in a second-order differential equation for displacement x as a function of time.

$$m\ddot{x} - c\dot{x} - \sum k\vec{l} = 0 \tag{2}$$

Mass positions are computed at each frame using Verlet integration.

The process is divided into three steps described in Figure 1. First an image is built from a 3D scene through framebuffer to texture rendering and mapped to a quad for image analysis (quad Q in figure 2.a & 2.b). The 3D scene is a video textured quad in case of pre-recorded or live video animation (V in figure 2.a). In case of MSS animation through direct manipulation of geometrical objects, any 3D scene with interactable or draggable elements can be used (S in figure 2.b). Both live video and interactive 3D scene allow for gesture-based animation. The framebuffer textured quad Q is rendered through shaders that can perform the desired real-time special effects. The output image is read back to the CPU and a luminance histogram is built. Each pixel is stored in the bucket associated with its luminance level.

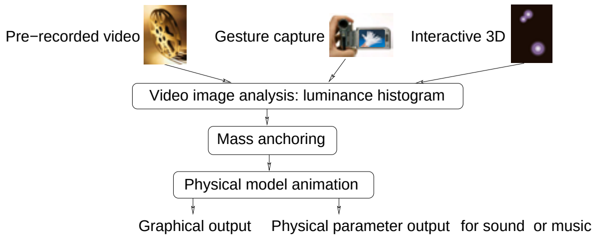


Figure 1: *Allegra* architecture

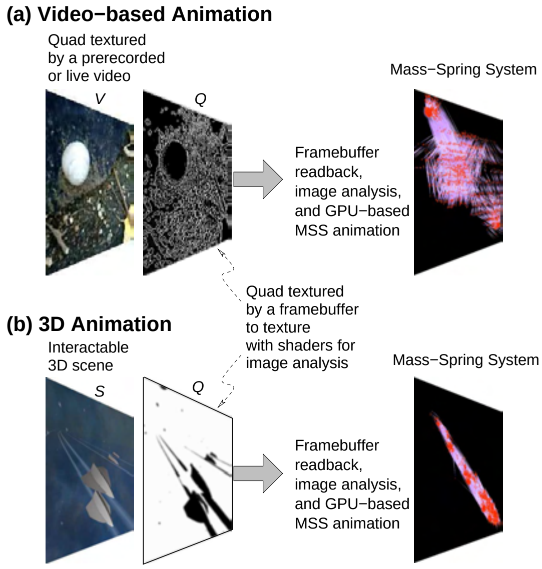


Figure 2: Graphical scene for video- or 3D-based MSS animation

In a second step, the brightest pixels are used to anchor one or both ends of each mass-spring chain to the image. The luminance threshold is defined by the histogram and the number of chain ends to allocate. We apply a conservative approach to pixel allocation so that anchors remain as stable as possible from one frame to the next one: only pixels that have shifted from a value above the threshold to a value below the threshold are deallocated and replaced by pixels above the threshold.

The third and last step is a multipass GPU-based MSS animation. As in [6], the MSS dynamics is computed in four steps: (1) edge-based force computation, (2) right and left force summation, (3) Verlet integration, and (4) particle displacement and graphical rendering. The MSS topology (chains of springs connecting two masses) is transmitted only once to the GPU at the first frame. The parameters (elasticity, rest length, damping, masses...) are updated dynamically and modify the MSS behavior. Masses and springs are rendered as translucent quads parallel to the projection plane. Since the mass and spring speeds and locations are stored in Frame Buffer Objects, they can be rendered to textures that are read back in the CPU memory. These values can then be used to parameterize an external application such as a music synthesizer.

Allegra is implemented with *Virtual Choreographer* (VirChor)[1] a generic interactive 3D-engine based on an XML representation of multimedia data and behaviors. Video streaming relies on *ffmpeg*[2] and the shading language for real-time image processing is *Cg*[3].

4. Three Video and Control Setups

In *Allegra*, the dynamics and the graphics of the visual rendering depend on 3 factors. (1) **The video source**: The video motion has a great impact on rendering. In the case of slow video motions, the rendering is smooth and elastic, whereas quickly moving targets produce ray effects. (2) **The MSS parameters**: The impact of MSS physical characteristics on their behavior is well known of all the researchers and artists who have worked in the domain of physically controlled graphics. An increase of damping coefficient stretches the springs and slows down the mass motion as in a viscous liquid whereas mass values and elasticity impact the resonance of the system. (3) **The visual aspects of masses and springs**: Textures and image processing techniques such as blur, echo, edge detection, dithering... are other means to control the visual rendering.

The video source determines the type of interaction. In the case of live video, the brightest parts of a performer's body can be used to control the location of the chain ends and thus directly manipulate the MSS. Prerecorded video is used for non-interactive animation, even though the user can interactively tune the animation parameters to modify the graphical rendering.

4.1. Direct Manipulation through 3D Scenes

Of the two possibilities of direct manipulation through video analysis, we first consider the case of a 3D scene in which geometrical elements can be dragged, and changed in their color and aspect. A wide range of geometrical scenes can be designed, we only consider here a scene made of two sprites (figure 3.a).

Without any motion of the sprites, all the springs and masses are grouped around these two sprites (figure 3.b). By dragging the masses around the display, the trailing mass-spring chains build vanishing streaks (figure 3.c). A different shape of the geometrical objects, such as long segments instead of disks, results in very different dynamic patterns (figure 3.d).

Such an interactive graphics can also be projected on digital tabletops such as MERL's *UbiTable*[4], Jeff Han's multitouch table[5], or Microsoft's *Surface*™[6]. With such devices, the dragging of the 2D or 3D can be easily controlled by the location of the fingers on the table. Additional interactors can be offered to let the users modify dynamically the shape and orientation of the sprites. The output of such an instrument can be seen as a new type of digital painting that evolves over time.

[1] `http://virchor.sf.net`
[2] `http://ffmpeg.mplayerhq.hu`
[3] `http://developer.nvidia.com`
[4] `http://www.merl.com/projects/UbiTable/`

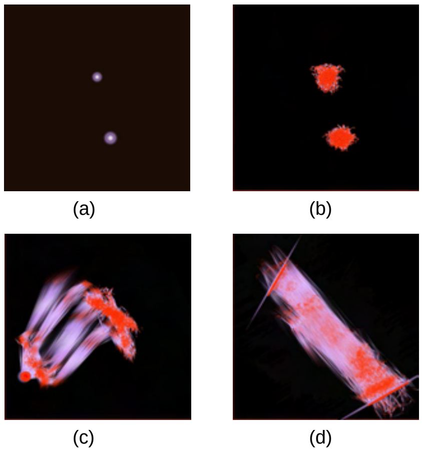


Figure 3: Direct manipulation with 2 draggable sprites

4.2. Bimanual Control through Live Video

The second possibility for MSS direct manipulation relies on the user's video capture (figure 4.a). The feeling for the user is different from the preceding case because the masses are anchored to a bright part of the body. In the setup proposed here, the user is dressed in black on a black background and her/his hands make the bright part of the image.

In the case of long mass-spring chains (10 to 15 springs), the MSS behaves as an elastic matter that sticks to the hands (figures 4.b & 4.c). Shorter chains and less damping result in very different behaviors that look like rays of light or springing water (figure 4.d).

4.3. Special Effects for Prerecorded Video

The animation of the physical models can also be controlled by pre-recorded videos that do not require human involvement at runtime. Such a system is a generator of special effects that matches the animation of the video image (figures 5.a & 5.b). The interaction, in this case, is more focused on the dynamic parameterization of the model than on its animation.

The graphical aspect of the output can also be modified by changing the shape of the masses: small circles in figure 5.c and words in figure 5.d.

[5] `http://cs.nyu.edu/~jhan/ftirtouch/`
[6] `http://www.microsoft.com/surface/`

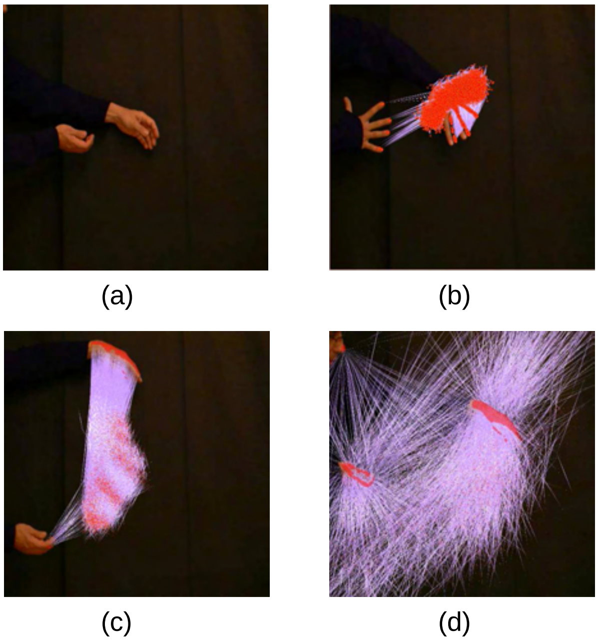


Figure 4: Bimanual interaction

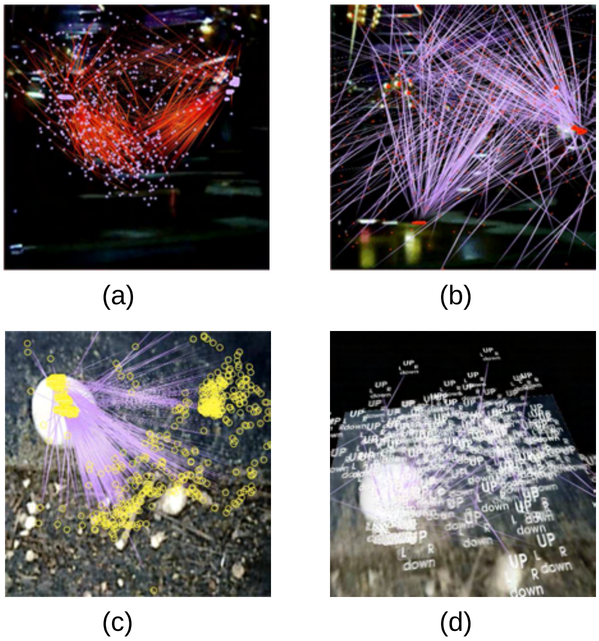


Figure 5: MSS animation from prerecorded video (video film is from Bertrand Planes)

5. Perspectives

Allegra can be seen as an end in itself or just as a component in a larger environment. As we have shown, the input of the system is basically any dynamic image. The input of the system could be its output at the preceding frame, and thus result is some kind of graphical feedback loop and self resonance as already studied in [8]. An immersive spatialized environment, such as a CAVE™ can also be used to place the user inside the particle flow as in [3]. Last the output values of the mass speeds and locations can be used as input for multi-sense rendering by controlling haptic or sonic outputs.

References

[1] W. Baxter, Y. Liu, and M. C. Lin. A viscous paint model for interactive applications. *Comput. Animat. Virtual Worlds*, 15(3-4):433–441, 2004.

[2] A. Camurri, G. Volpe, G. D. Poli, and M. Leman. Communicating expressiveness and affect in multimodal interactive systems. *IEEE MultiMedia*, 12(1):43–53, 2005.

[3] T. Chen, S. Fels, and T. Schiphorst. Flowfield: investigating the semantics of caress. In *Proceedings of SIGGRAPH '02*, pages 185–185. ACM Press, 2002.

[4] T. Cockshott. *Wet and Sticky: A novel model for computer based painting*. PhD thesis, University of Glasgow, 1991.

[5] J.-L. Florens, C. Cadoz, and A. Luciani. A real-time workstation for physical model of multi-sensorial and gesturally controlled instrument. In *Proceedings of the 1998 International Computer Music Conference*, pages 518–525. ICMA, 1998.

[6] J. Georgii and R. Westermann. Mass-spring systems on the GPU. *Simulation Modelling Practice and Theory*, 13:693–702, 2005.

[7] D. M. Howard and S. Rimell. Real-time gesture-controlled physical modelling music synthesis with tactile feedback. *Eurasip Journal On Applied Signal Processing*, 7:1001–1006, 2004.

[8] C. Jacquemin and S. de Laubier. Transmodal feedback as a new perspective for audio-visual effects. In *NIME 2006 International Conference on New Interfaces for Musical Expression*, pages 156–161, Paris, France, May 2006. NIME.

[9] A. Momeni and C. Henry. Dynamic independent mapping layers for concurrent control of audio and video synthesis. *Computer Music Journal*, 30(1):49–66, 2006.

[10] J. Stam. Stable fluids. In *Proceedings of SIGGRAPH '99*, pages 121–128. ACM Press/Addison-Wesley Publishing Co., 1999.

[11] K. A. Woolford and C. Guedes. Particulate matters: generating particle flows from human movement. In *MULTIMEDIA '07: Proceedings of ACM Multimedia 2007*, pages 691–696. ACM Press, 2007.

A Summary of Formats for Streaming and Storing Music-Related Movement and Gesture data

Alexander Refsum Jensenius,[a] Nicolas Castagné,[b] Antonio Camurri,[c]
Esteban Maestre,[d] Joseph Malloch,[e] Douglas McGilvray [f]

[a]University of Oslo, Musical Gestures Group, a.r.jensenius@imv.uio.no
[b] ACROE, Grenoble, nicolas.castagne@imag.fr
[c] University of Genoa, Infomus lab DIST, antonio.camurri@unige.it
[d] Pompeu Fabra University, Music Technology Group, emaestre@iua.upf.edu
[e] McGill University, IDMIL, CIRMMT, joseph.malloch@mcgill.ca
[f] University of Glasgow, Centre for Music Technology, d.mcgilvray@elec.gla.ac.uk

Abstract

This paper summarises a panel discussion at the 2007 International Computer Music Conference on movement and gesture data formats, presents some of the formats currently in development in the computer music community, and outlines some of the challenges involved in future development.

1 Introduction

The rapid growth in research on enactive interfaces over the last years, and on movement and gesture in general, have shown the need for better methods, tools and techniques for handling what we will here refer to as *movement and gesture data*. One important challenge is the lack of generic formats for handling such data, something which often leads to compatibility problems when working with various hardware and software solutions. This issue has emerged as an important research topic in the *computer music* community over the last years. Considering that a computer music point of view may stimulate a larger discussion in the Enactive audience, this paper provides an overview of the solutions that are currently being worked on in this field.

While we have formats and standards for handling audio (AIFF, MPEG, etc.), audio analysis (SDIF), video (MPEG, QT, etc.), music notation (MusicXML), musical control data (OSC), etc., there are no widespread formats, nor structured approaches, for handling music-related movement and gesture data. In fact, most researchers store their data without using any specific format, or use the format of the specific device or application at hand [5]. This is a practical problem not only for the single researcher working with various types of equipment, but it also effectively hinders the sharing of data, tools and research methods between institutions.

Movement and gesture related studies have gained interest in the computer music community over the last years, and several research groups have started to work on solutions for standardising the way we store and stream movement and gesture data. Since several of these initiatives seemed to be unknown in the computer music community, we invited a number of researchers involved in the development of various movement and gesture data formats to a panel discussion at the 2007 International Computer Music Conference[1] in Copenhagen, Denmark [2]. This paper summarises the panel discussion, provides an overview of some of the formats in development by the authors, and points out some challenges for future development. Focusing mainly on the point of view of the computer music community, these formats may also be a starting point for a wider approach to encoding movement and gesture data.

2 Structuring Low level signals

A major challenge in the development of formats for handling movement and gesture data seems to be the lack of defining and structuring low-level movement and gesture *signals* or *streams*. We deal with low-level data representing performed movements and gestures, but there is no common agreement on how to describe, structure and encode such low-level data. While it is sometimes possible to work with device specific data, there is a growing need to record, store and exchange low-level data in a more generic way.

In the same way as the PCM audio format served as a foundation for the development of research on audio, we

[1] http://www.icmc2007.net

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

find that the establishment of a generic, minimal format to structure and encode low level movement and gesture signals is crucial for further research on movement and gesture in fields such as computer music, enactive interfaces, computer graphics, virtual reality, etc.

2.1 Motion Capture Formats

The motion capture community has introduced a number of formats dedicated to storing and structuring motion capture data over the years. Some of these formats are used in the computer music community, but they are often far from sufficient for many of our needs, and therefore often create more problems than they solve.

One problem with several of these formats is that they are proprietary and designed to accompany specific hardware, something which does not give the openness and expandability that we look for. Another problem is that many of these motion capture formats focus on full-body motion-capture streams based on an articulated skeleton and a 3D-representation. This is often not general enough for many computer music applications where we are not only interested in describing human bodies, but also devices with different morphologies and dimensions, as well as information about tactility and haptics in the devices. In general, we therefore find existing motion capture formats too specific when it comes to dimensions, structure, number of degrees of freedom, and frequency characteristics (often also limited by storing in ASCII-files).

Yet another problem with many of the motion capture formats is the lack of possibilities for synchronising low-level data with mid- and high-level analytical results, as well as other types of data (e.g. music notation) and media (e.g. audio and video). This calls for more generic formats that can synchronise data with various resolutions and sampling rates (see section 3).

2.2 GMS

The Gesture and Motion Signal (GMS) format[2] has been developed in the EU Enactive Network of Excellence[3] by a subgroup of partners headed by the ACROE group [4]. It is a binary format intended for structuring, storing and streaming low-level movement and gesture signals as generically as possible, not only for computer music applications.

In GMS, a *gesture scene* can be encoded at any frequency rate (e.g. 100 Hz to a few tens KHz) and it is based on a two-level structure made of *gesture channels* and *gesture units* (Figure 1). A gesture channel allows for structuring the dimensions of the performed gesture; it can correspond either to an intensive variable (e.g. position) or extensive variable (e.g. force), that can be either 1D, 2D or 3D. A gesture unit is made of a group of channels, and allows for structuring various recorded points/forces in a meaningful manner.

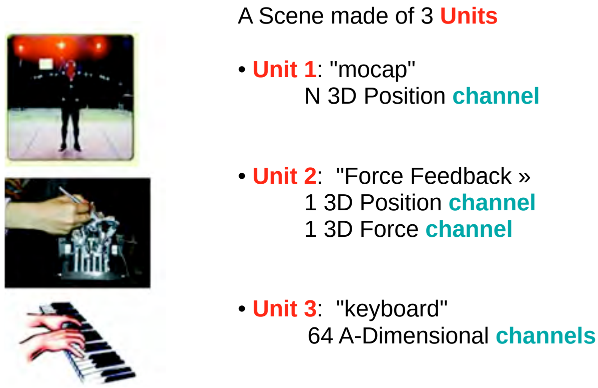


Figure 1: An example of a gesture scene structured and encoded with GMS

3 Mid- and high-level data

Much of the analysis and usage of music-related movement and gesture data is happening at what may be called mid- and high levels, e.g. focusing on phrases, expressivity, emotional response, etc. For this reason we need to find solutions to handle such data in a structured manner and to synchronise such data with the low-level data they are often derived from. There are several research groups involved in finding solutions for handling the structuring of such data, and three formats are currently being developed: GDIF, PML, XMI.

3.1 GDIF

The development of the Gesture Description Interchange Format (GDIF)[4] is a collaborative effort between researchers at the University of Oslo, McGill University and Pompeu Fabra university [3]. The focus is on creating structures for handling different levels of movement and gesture data: from raw data to higher level descriptors, as well as secure synchronisation with other types of data and media.

GDIF development is mainly focused on *what* to store and not *how* to store it, and is therefore based on existing formats and protocols, e.g. OSC, SDIF and XML (Figure 2). This allows for both streaming and storage, as well as compatibility with various computer music software and hardware. For realtime control, GDIF has been tested to control spatialisation [8] and creating a more structured and flexible approach to setting up mappings between various sensor devices and sound engines [7]. For the analysis of musical gestures,

[2] http://acroe.imag.fr/gms/
[3] http://www.enactivenetwork.org
[4] http://musicalgestures.uio.no

an XML-based implementation of GDIF is being developed for creating performance databases, exemplified through violin performance in [6].

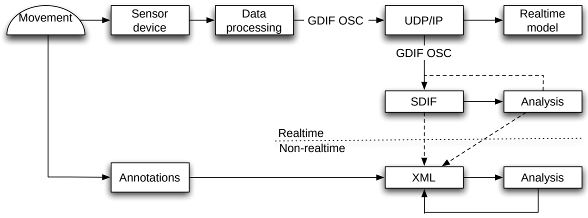


Figure 2: A GDIF setup for handling both streaming and storage of data using OSC, SDIF and XML

3.2 PML

Performance Markup Language (PML)[5] is an XML-based representation intended to facilitate the investigation of issues relating to musical performance. To investigate these issues, it is necessary to analyse performance artefacts in the context of the score. Therefore, the basic content of a PML file comprises the score, a basic markup of performance events, and a description of the correlation between individual performance and score objects.

The strict hierarchy of XML is not naturally suited to the multiple overlapping structures which are required to adequately describe musical information. Therefore, PML encourages information to be stored within separate hierarchies. These hierarchies can cross reference information in other informational hierarchies using internal relational links and pointers to locations within external files in formats such as PCM audio or GMS low level signals. Therefore PML allows existing formats to be combined into one representational system, allowing existing tools for manipulation of score, audio, video and gesture to be used.

3.3 EyesWeb XMI

The new EyesWeb XMI[6] (eXtended Multimodal Interaction) proposes a multi-layered gesture processing framework containing three layers: *MoCap*, *trajectories*, *cues and gestures*. This allows for working with: i) real-time, multimodal processing and interactive systems, and ii) data analysis and synchronised processing of pre-recorded data (see for example [1]).

EyesWeb XMI is supporting various geometric data types and compound data types such as collections to face the multifaceted and multi-layered problems that arise in gesture analysis research. It is possible to represent and process point-light display data, multicamera and multisensor data, as well as collections of different data and expressive gesture cues. A full set of automatic converters between the different layers and data types issupported.

[5] http://www.n-ism.org/Projects/pml.php

[6] http://www.eyesweb.org

4 Other Formats and Protocols

A critical aspect in computer music is the need to synchronise data with other types of data and media. Besides formats for handling audio, video, images and notation, there are two formats that have been established as ”standards” over the last decade: SDIF and OSC.

4.1 SDIF

The Sound Description Interchange Format (SDIF) was originally developed for handling audio and audio analysis data [10], and has been implemented in a number of software and programming environments [9]. The SDIF specification and implementation has already tackled a number of challenges relating to synchronisation of multiple streams of exogenous data, including high-speed data streams.

Even though SDIF was originally developed for storing audio data, it is a ”container” format that could easily be extended to carry necessary low-, mid- or high-level movement and gesture data (section 2 and 3). This, however, still requires development of taxonomies and structures for such data, as currently being developed in GMS, GDIF and PML.

4.2 OSC

Open Sound Control[7] (OSC) is an open, transport-independent, message-based protocol for communication between music hardware and software systems [12]. OSC has received increased interest over the years and is currently the *de facto* communication standard in the computer music research community, and is also slowly being introduced in various commercial systems (as an alternative to MIDI).

OSC does not solve the encoding and structuring of movement and gesture data, only the transport of the data. That is why it is necessary to develop solutions for a structured approach to creating OSC namespaces for streaming movement and gesture data, such as GDIF.

5 Summary

Standards-making seems to be an ongoing, iterative activity in the computer music community [11], and one can argue that the most successful formats in use are the ones that started by solving a specific problem for later

[7] http://www.opensoundcontrol.org

to be developed into a more generic standard. Such a bottom-up approach is, indeed, the approach taken by several of the authors in their various developmental efforts, including GDIF, GMS, PML and EyesWeb XMI.

The panel discussion at ICMC, and this follow-up paper has presented some of the current challenges and research efforts when it comes to movement and gesture data formats in the field of computer music. Some key elements for future development are to:

- create solutions for both performance (streaming) and analysis (storage).
- define, structure and encode low-level continuous movement and gesture signal data with different frequencies, resolutions, dimensions, etc., including various feedback loops.
- define, structure and encode mid- and high-level analytical data and descriptors, and synchronise these with related low-level data.
- handle synchronisation with musical notation, other types of data (e.g. annotations) and media (audio and video)
- support already existing formats and protocols used in the community, e.g. SDIF, OSC, MusicXML.

The various formats developed by the authors approach different aspects of the above-mentioned problems, and by uniting research efforts it may be possible to ensure interoperability between the different formats. An important point here is that of cross-disciplinary collaboration. Similar problems relating to structuring and encoding movement and gesture data are currently being tackled by researchers in various fields, and much research still needs to be carried out both conceptually and technologically. By joining efforts, we may be able to more efficiently reach our goals of generic solutions for handling movement and gesture data. Hopefully this paper may stimulate such further collaboration.

Acknowledgements

Thanks to Diemo Schwarz, Matt Wright, Stuart Pullinger and Ben Knapp for participating in the ICMC panel, and for valuable comments and suggestions.

References

[1] A. Camurri, G. Castellano, R. Cowie, D. Glowinski, B. Knapp, C. L. Krumhansl, O. Villon, and G. Volpe. The premio paganini project: a multimodal gesture-based approach for explaining emotional processes in music performance. In *Gesture Workshop, Lisbon*, 2006.

[2] A. R. Jensenius, A. Camurri, N. Castagne, E. Maestre, J. Malloch, D. McGilvray, D. Schwarz, and M. Wright. Panel: the need of formats for streaming and storing music-related movement and gesture data. In *Proceedings of the 2007 International Computer Music Conference*, Copenhagen, Denmark, forthcoming 2007. San Francisco: ICMA.

[3] A. R. Jensenius, T. Kvifte, and R. I. Godøy. Towards a gesture description interchange format. In *NIME '06: Proceedings of the 2006 International Conference on New Interfaces for Musical Expression*, pages 176–179, Paris, France, 2006. Paris: IRCAM – Centre Pompidou.

[4] A. Luciani, M. Evrard, N. Castagné, D. Couroussé, J.-L. Florens, and C. Cadoz. A basic gesture and motion format for virtual reality multisensory applications. In *Proceedings of the 1st international Conference on Computer Graphics Theory and Applications, Setubal, Portugal, March 2006*, Setubal, Portugal, 2006.

[5] A. Luciani, M. Evrard, D. Courousse, N. Castagne, I. Summers, A. Brady, P. Villella, F. Salsedo, O. Portillo, C. A. Avizzano, M. Raspolli, M. Bergamasco, G. Volpe, B. Mazzarino, M. M. Wanderley, A. R. Jensenius, R. I. Godøy, B. Bardy, T. Stoffregen, G. De Poli, A. Degotzen, F. Avanzini, A. Roda, L. Mion, D'Inca, C. Trestino, and D. Pirro. Report on Gesture Format. State of the Art. Partners' propositions. Deliverable 1 D.RD3.3.1, IST-2004-002114-ENACTIVE Network of Excellence, 2006.

[6] E. Maestre, J. Janer, M. Blaauw, A. Pérez, and E. Guaus. Acquisition of violin instrumental gestures using a commercial EMF tracking device. In *Proceedings of the 2007 International Computer Music Conference*, Copenhagen, Denmark, 2007. San Francisco: ICMA.

[7] J. Malloch, S. Sinclair, and M. M. Wanderley. From controller to sound: Tools for collaborative development of digital musical instruments. In *Proceedings of the 2007 International Computer Music Conference*, Copenhagen, Denmark, 2007. San Francisco: ICMA.

[8] M. T. Marshall, N. Peters, A. R. Jensenius, J. Boissinot, M. M. Wanderley, and J. Braasch. On the development of a system for gesture control of spatialization. In *Proceedings of the International Computer Music Conference*, pages 360–366, New Orleans, LA, 2006. San Francisco: ICMA.

[9] D. Schwarz and M. Wright. Extensions and applications of the SDIF sound description interchange format. In *Proceedings of the 2000 International Computer Music Conference*, pages 481–484, Berlin, Germany, 2000. San Francisco: ICMA.

[10] M. Wright, A. Chaudhary, A. Freed, D. Wessel, X. Rodet, D. Virolle, R. Woehrmann, and X. Serra. New applications of the sound description interchange format. In *Proceedings of the 1998 International Computer Music Conference*, pages 276–279, Ann Arbor, MI, 1998. San Francisco: ICMA.

[11] M. Wright, R. Dannenberg, S. Pope, X. Rodet, X. Serra, and D. Wessel. Panel: Standards from the computer music community. In *Proceedings of the 2004 International Computer Music Conference, Miami, FL*, pages 711–714, 2004.

[12] M. Wright, A. Freed, and A. Momeni. Opensound control: State of the art 2003. In *NIME '03: Proceedings of the 2003 International Conference on New Interfaces for Musical Expression*, 2003.

Towards an Enactive Epistemology of Technics

Armen Khatchatourov John Stewart Charles Lenay
COSTECH – UTC, France
E-mail: armen.khatchatourov@utc.fr, john.stewart@utc.fr, charles.lenay@utc.fr

Abstract

The aim of this paper is to introduce an epistemological framework in which the technics are considered from the very start in their intertwining with the human. We show in which sense the enactive paradigm can provide the basis for such an epistemology.

Our thesis is that all technical artifacts, from stone tools to cars to computers, are "enactive interfaces" that mediate the structural coupling between human beings and the world they live in, and hence bring forth a particular world of lived experience.

The social dimension of this approach to technics is also discussed.

1. Introduction

The basic scheme for considering enaction is the dynamic sensory-motor coupling between an organism and its environment.

The sensory inputs, S, are used to guide the actions A; the actions A modify the environment and/or the relation of the organism to its environment, and hence modify in return the sensory inputs. This basic scheme applies to all living beings. In the 1920's von Uexküll characterized "animal worlds" on the basis of sensori-motor contingencies as they function in ecological context.

What the world "is" for the organism amounts to neither more nor less than the consequences of its actions for its sensory inputs; or to its "sensori-motor contingencies" [1]; and this in turn clearly depends on the repertoire of possible actions. Without action, there is no "world" and no perception.

There is a deep affinity between this approach, the "enactive" approach of Varela [2], and the ecological psychology [3] according to which perception is not a matter of computational representation, but rather a "direct" perception of "affordances", i.e. potential actions as such. This affinity lies, as we understand it, in (a) a non-representationalist framework, and (b) in the fact that "rules" or "laws of control" [3] or "contingencies" [2] are not pre-given but emerge from the interaction between an organism and its environment. For the purposes of this article, we are not going any further in the description of (discrepancies between) these approaches, we rather propose to consider what is specific to human beings.

One of the major characteristics of "human worlds" is that the sensory-motor coupling is mediated by technical artifacts, this leading to two radical innovations (Fig. 1)

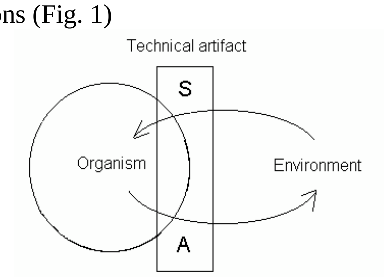


Figure 1. Mediated sensorimotor coupling.

Firstly, the range of possible sensory inputs and the repertoire of possible actions are greatly increased, without any limits other than the invention and fabrication of new artifacts. This is clear for the new possibilities of action which are created by tools, from hammers to power-tools. It is also clear for instruments - microscopes, telescopes, radios and so on resulting in sensory inputs which are strictly impossible without the devices in question.

More generally, but less obviously, technical artifacts organize sensory experience: think of the world of skier, which is impossible without the artifact. Even when we are not actually skiing, our perception of the mountain is determined by the possibility (i.e. virtual action) of skiing and the correlative sensations. So this first point can be understood more profoundly: in case of contemporary humans, there hardly are any "natural" perceptions or relations to the world: our sensory-motor coupling is always fashioned, at least virtually, by technical artifacts. [4]

Secondly, technical artifacts are not irremediably fixed to the body. More precisely, technical artifacts exist in two "modes": "in hand" and "put down". When a technical artifact is "in hand", being used, it becomes a

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19^{th}-22^{nd}, 2007

prosthetic extension of the body; correlatively, the artifact disappears from consciousness, and the attention of the human subject is focused on the "world" that comes about (think again of the "world of the skier", for example). Artifacts, like the body, are normally "transparent" to the subject; as Heidegger [5] has pointed out, they are only noticed when they are dysfunctional (a wobbly hammer or a twisted ankle). However, unlike biological organs, technical artifacts can also be "put down": separated from the body, they can now become objects of attention. In this mode, their objective physical proprieties can be perceived; they can be invented, fabricated, repaired and so on [6]. The whole question of learning can be seen as the back-and-forth movement between these two modes. This explains also the radical innovative potential of technical artifacts.

2. A categorization of technical artifacts

Before going further, it is useful to give a more complete categorization of technical artifacts, which can be roughly divided into 3 types. The first type we discussed above can be called "extensions of the body": tools and sensory instruments. But there is also a second type of artifact, consisting of deliberate modifications of the environment: roads, buildings, fields and so on. It is even more obvious that this second type of artifact also modifies the world that human beings live in.

A third sort can be called "semiotic artifacts". Here, the "actions" consist in emitting signals, and the sensory input is specifically geared to the reception of these signals. If the conditions that trigger the emission of a signal and the response of the receiver are appropriate, this leads to a co-ordination of actions, and constitutes the basic form of communication, present already in animal world. The human inventions are: first of all, language itself; and then multiple technical inventions: writing, printing, and computers. It is worth noting that the computers are not only semiotic artifacts, but also sensori-motor devices which comprise a certain repertoire of real actions (moving mouse, joysticks, etc.) with, in return, an increasing range of sensory inputs (visual patterns, sounds etc.); regularities are established between action and sensation in this case as for the first type of artifacts.

This categorization can be useful for analytical purposes; but it is important to note that in practice, technical artifacts do not function in isolation from each other, but form technical systems with a synergy between these three types. For example, roads (type 2) go together with cars (type 1), their synergy being organized by maps and plans (type 3). A possible use of the term "technology" (techno-logos) is to designate the situation where there is linguistic communication about the design, fabrication and use of technical artifacts.

Following [5], some precisions can also be brought to the distinction between the two modes we mentioned above. The difference between in hand and put-down is not simply between attached/not attached to the body.

First f all, there are two relatively independent levels of dividing:

1. Between put-down and in-hand: put-down corresponds to the mode in which the artifact is the object of the explicit attention as an assembly of the matter with certain proprieties (the specifically scientific mode of relation to the object). One can think on the difference between designing and riding the bicycle. The in-hand mode is the mode in which the user is engaged in the activity, and in which, under normal conditions, the artifact is transparent, one feels it like the extension of the body, not like the object of the physics.

2. Between a normally functioning and a broken artifact. Now comes the situation in which the artifact is broken. In this situation, the artifact switches from in-hand to put-down: instead of riding the bicycle and being engaged in the sensory-motor activity, one examines the broken chain as something having being made of the material with bad resistance, etc. So one can as she wish be in different attitudes to artifact: consider it as in-hand or put-down (when maintaining the technical device, one puts it in the put-down mode). But the situation when the artifact breaks is particular, because it forces the user to consider it as put-down.

It is the same case with the computers, even in the Virtual Reality. As a user, one does not care about what is going in the computer, it becomes a transparent "equipment". When the artifact is broken, the user will check cables, electricity, she will consider it as an object of science and technology, and the artifact is not a transparent mean of action anymore.

The difficulty comes when we consider the fact that in the put-down mode, the designer is also engaged in the activity. But in a different way: the artifact is not a mean of action. In fact, when one is maintaining/designing or doing scientific research, she is using other artifacts (pencils/CAD/hammer or measurement instruments), which are in-hand as means of action, and which are transparent to the user. Thus one can see the put-down mode as a derivative from the most fundamental in-hand mode.

Now the in-hand mode was provisionally defined as an attachment to the body, in order to underline the fact that it is transparent and fits into action. But in fact the artifact can be not attached to the body, but still in-hand. The road for example is not attached to the body, but is still in-hand as a transparent mean of action. Being on the road, one does not consider the road as the physical proprieties of tarmac in the way the science/technology do, but rather as a possibility to get where she wants to; the lighting pole on the road is not attached to the body, but it is still in-hand because it is also a mean of action of going there; moreover, the light coming a certain way, one takes it into account without explicitly thinking on its properties, and adapt her sensory-motor activity when riding a bicycle

It stands to reason that there is still a difference between the artifact that are actually attached to the body, and which are not, but the first level of distinction seems to be between in-hand (in a broad sense) and put-down. In this broad sense, the artifacts are in in-hand mode when they (a) fit into action, (b) change sensory-motor loops, (c) are transparent, i.e. not explicitly noticed, disappear from consciousness in aid of the world they bring forth.

3. The social dimension

The fact that technical artifacts exist in the mode of being "put down" has an important consequence: the persons who design and make technical artifacts are, generally, not the same as those who use them. Thus, technological development goes together with a division of labour and, correlatively, the development of mechanisms of social synthesis (exchange, market economies) which organize the integration of technical systems as functional wholes.

Traditionally, the technology is usually considered as a "black box", as intrinsically neutral means to pre-defined ends. The approach outlined here leads to a new perspective in which technology occupies a central position. The work of engineers has immense social significance because, in fine, the choices of technological devices fashion the human condition itself, by manufacturing interfaces that change the means of action, and influence sensations.

This introduces the debate about the usage of the artifacts, which is of particular importance in the field of Human-Computer Interaction where the design of interfaces is supposed to be adapted to the user. The question is then to articulate the design with the user's needs, abilities and knowledge, and to build possible "enactive" interfaces. This raises multiple epistemological questions: Is the knowledge of the usage situated in the user as an acquired sensorimotor knowledge? Does the quality of the interaction depend only on this user's knowledge? How to build an artifact responding to the sensorimotor knowledge of the user?

Since the sensorimotor knowledge is not something independent from the practice of artifacts, it seems difficult to say that it is situated in the user. If the artifact modifies the established sensory-motor contingencies, then the enactive knowledge depends on the artifacts. In other words, for human beings, it seems impossible to talk about a standalone user, on whose knowledge depends the use of the artifact, and the ability to make it enactive. From our point of view, we need to understand how the enaction takes place between the two terms, the user and the artifact.

Still, "enactive" is a quality that does relate to the individual, and the experience of an enacted world as a world of possibilities is always for a human (who is always technically equipped, even if she doesn't actually use any interface), and the artifact alone does not enact anything. But if the capacity to enact lies in the user, artifacts do always change the quality of enaction and human's experience.

So already for a "single" user the enactive knowledge is something situated "between" the user and the artifact, but what about the social exposure? The couple "artifact / sensory-motor contingencies" is something that does evolve on the scale of the society, and the problem of usage is something intrinsically social. That's why it is difficult to report this problem to enactive knowledge of a single user.

What is enactive, it is not the interface itself, neither the usage alone, it is the "combination" of them. If one designs a very "enactive" interface, but there is no social acceptance or implication, in the best case the usage will be restricted to a narrow community. But the contrary is also true: if the interface is not appropriated, there will be no enaction (in the following sense: no "good quality" of relation between the human and the world) even if there is a wide social exposure. So, we need to distinguish two sorts of Enactive Interfaces: in a broad sense, every technical artifact is enactive because it does modify the sensory-motor contingencies, and bring forth a particular lived experience, even if the artifact is really "constraining"; in a strong sense, the criteria for the interface to be enactive (good "quality" of interaction, transparency, etc) are actually still to find.

But this is probably not enough. If we continue to think, - and that was the mainstream of industrial engineers -, that it is sufficient to design an interface that seems good to designers, we would be probably wrong. Many works on the anthropology of usage and on involving the end-users in the process of design seem to go in this direction.

Moreover, what one accepts as a quality of interaction, is not something independent on technology itself, more precisely on the socially accepted aspect of technology or, let's say, its historical aspect. (It is not sure that today's cameraphones are really useful and enactive interfaces, they are however widely socially accepted as something having a quality of interaction). In other words, the artifacts are not only responding to functional criteria, they are also, as Leroi-Gourhan [7] for example has pointed out, a support of figurative aesthetics, and this may be to the detriment of the pure functionality. This could help us to understand in which way the acceptance of the artifacts is related to sensory-motor knowledge: this knowledge is always socially and technically transmitted and determined. However, it is worth noting that in any case we are not talking about a technological determinism: the question is how the social structures "arrange" with the technology, and not what technology "imposes" by itself. The core question is that it is difficult to know which interfaces will have the social implications.

Would the artifact have or not the social exposure is not something lying in the technology if one considers the technology as the pure functionality of the artifact; but it is something lying in the technology if one considers the technology also as something intrinsically socially constructed, and also if one considers the social structures (for example the exposure of the artifact related to the socially accepted criteria of aesthetics) as something technically transmitted.

4. Discussion

The preceding considerations may not seem particularly controversial, but they have some controversial consequences. The term "Interface" is clearly of central importance. However, the term itself is the vehicle of an ambiguity that requires clarification: "interface" between what and what?

As we understand it, the term "interface" is properly used as the interface between an organism (human or otherwise) and its environment. Thus, the basic "interfaces" are the biological sensory and motor organs; for humans, technical artifacts are extensions to these basic interfaces, but they remain interfaces. New technical devices constitute new "worlds": think for example of the "world of the skier". But note this: we do not talk about the "interface" between the man and the ski; the ski is the interface between the man and the snowy mountain, or better still between the skier and the "skiing world" that is brought forth.

Does this change in the case of computers? Our point of view is that computers are basically technical devices, and should be treated in the same way as other technical devices. Certainly, they are devices of a special sort, and the "worlds" that are brought forth when a human being uses them are a special sort of "world"; but the interaction that occurs (that is mediated by the machine) is between the human being and this "world"; it is not an interaction between the human being and the machine. Thus, there is something deeply wrong in the very phrase "Human-Computer Interface". Of course, "HCI" has become a hackneyed term, but this engrained (mis)-use does not make it correct. The basic problem lies in the implication that human beings and computers are entities of the same sort, so that they could "interact" on a basis of equality. This would only be correct if one whole-heartedly embraces the representational paradigm according to which humans function like computers; but as we understand it, the enactive approach rejects this classical paradigm in cognitive science.

Finally, an interesting question that arises is the status of "virtual reality". In this case, it does seem as though the computer is playing the role of "the world", by providing the sensory consequences of actions on the part of the human being. But even here, note that the experience of a human being immersed in a "virtual reality" is not that of interacting with a computer; the human interacts with the entities that populate the "world" that has been brought about. We only become conscious of the computer (the interface) when a malfunction triggers the switch to the put-down mode; in normal functioning (the in-hand mode) the computer-interface disappears from consciousness.

This remark is in no way meant to decry the interest of "virtual realities"; on the contrary, such experiments are deeply revealing. What they show is that in order to create a "virtual reality", it is neither necessary nor sufficient to compute (in all its gory detail) the total physical reality – an impossible task anyway, as shown by flight simulators that have to fall back on analog models; what is required is neither more nor less than to provide the appropriate sensory returns to human actions. This helps, greatly, to bring home the point that what human beings experience in "natural" situations is not an objective world "as-it-is", but the sensory-motor contingencies of their embodied situation. Thus, interfaces and tools can permit (or not) humans to enact the world, and the world we live in depends on their design.

References

[1] O'Regan K.J. and Noë A., "A sensorimotor account of vision and visual consciousness." in *Behavioral and Brain Sciences* 24, 2001, pp. 5-115.

[2] Varela F., Thompson E. & Rosch E. The Embodied Mind. MIT Press, Boston, 1991.

[3] Warren, W. H. (1998). Visually controlled locomotion: 40 years later. *Ecological Psychology*, 10, 177-219.

[4] Khatchatourov, A et Auvray, M. "L'outil modifie-t-il la perception ou la rend-il possible?" in *Arob@se*, 2005, vol. 1. www.univ-rouen.fr/arobase

[5] Heidegger, M., Being and Time, State University of New York Press, 1996

[6] Lenay et al. Sensory Substitution, Limits and Perspectives, in Hatwell et al (eds), "Touching for Knowing", John Benjamins Publishers, Amesterdam, 2004

[7] Leroi-Gourhan, A., Gesture and Speech, MIT Press, 1993.

Devices for the Perception of Magnetic Fields

Ted Krueger
School of Architecture, Rensselaer Polytechnic Institute, USA
E-mail: krueger@rpi.edu

Abstract

Bach-y-Rita's 'sensory substitution' research was the first example of a class of perceptual prostheses that could, in principle, be designed using manufactured sensor technologies, processing circuitry, and appropriate transduction to the body. There are few examples of devices of this kind that are designed specifically to interface to environmental properties that are outside the normal range of human senses. This paper describes prototype devices that are intended to make magnetic fields directly perceivable.

1. Interfaces

Enactive interfaces are commonly understood as interfaces to physical objects: tools, instruments, machines, systems or representations. From my perspective within the field of architecture, the importance of objects is equaled, perhaps surpassed, by the importance of environments. The pervasiveness and diffusion of technologies and the inhabitation of increasingly synthetic domains suggests that interfaces to technological environments will gain importance. But intelligent, responsive and adaptive environments are only one part of a spectrum of environments that we inhabit – some highly technologized and some much less so. Can we develop interfaces only to the technological side of this spectrum? Maturana and Varela (1980) suggest that the organism (from the perspective of the organism) is informationally closed and this, in turn, requires the environment to be constructed out of information internal to it rather than received from the outside. This implies that the principles that underlie perception are applicable to the design of interfaces to technological objects, but also might be applied to interfacing to the environment itself. This paper is concerned with the design of devices that are interfaces to the environment of the wearer. They have as their objective enabling the perception of aspects of that environment that are not currently perceivable.

This paper describes some initial attempts to build devices that will result in the direct perception of magnetic fields. Direct perception means that the environmental property or dimension is available to immediate apprehension without the need for inferential mental processing (Stoffregan and Bardy 2001). Perceptions are not representations or depictions which have an entirely different experiential character (Stoffregan 2003). By this they are distinguished from the scientific visualization of data. The goal of my research is to generate experiences that are integrated with other percepts resulting in the awareness of additional environmental properties.

The discipline of architecture is commonly thought of as the design of objects such as buildings, systems, furnishings and the like or the configuration of environments. This understanding of architecture is consistent with an objectivist position. Its focus will be on material specification and metric description – the quantification of space. Another approach sees the task of architecture as the qualification of space in support of human activities. In this perspective, the activity of design is directly related to the process of perception and the construction of the realities within which we exist (Glanville 2006). Research into the process of perception and its testing and extension through the building of interfaces that bring an expanded awareness of the environment constitute fundamental research for the disciplines of design.

2. An Enactive Approach

In an 'enactive' view of perception, the dynamic sensori-motor coupling between an organism and its environment links changes in the sensory flux to the dynamics of the body and relies on the integration of afferent and efferent nerve activity. The organism, in this case, is active and exploratory; willfully engaging its environment, and by doing, so generating a portion of its sensory dynamics (Thompson and Varela 2001). It is the element of exploration that suggests that the objective of perception is the probing of conditions and the elaboration of the apprehended world rather than the development and maintenance of representations of an external objective reality. This direct linking between the activity of living and perception is in stark contrast to the abstraction into information and representation in the computationalist view.

The organism must be able process a vast array of nerve activity that has no *a priori* structure into patterns that become identified with the self and the external world. But in addition, objects in the world, their perceptual attributes and spatial locations must also be derived. In the same way, the organism has a spatial disposition in relation to the apprehended world and its own set of properties and potentials that must be realized. Phillipona, O'Regan and Nadal (2003) suggests that pattern recognition capabilities may be capable of extracting, in principle, invariants in the sensory flux from which self, world, objects and properties can be inferred. Knowledge is the patterns that we form.

O'Regan and Noë (2001) use the term 'sensorimotor contingency' to describe the systematic variations in sensor states that arise from the active engagement of an organism with its media. These contingencies give the various sense modalities their respective qualities. Vision is not the sense of light striking receptors in the retina as traditionally understood, but a mode of exploration mediated by distinctive sensorimotor relationships. Therefore, vision does not depend upon the eye or optic nerve. Sensory substitution strategies are cited as a case in point.

3. Interfacing the Environment

Beginning in the late 1960s, Bach-y-Rita developed a series of devices that linked the output of a video camera to an array of tactors applied to the surface of the forehead, back, arm, thigh, fingertip or tongue using vibrotactile or electrotactile stimulation (Bach-y-Rita 1972, Bach-y-Rita *et al* 2003). Significantly, the apprehension of these spatial conditions only arises with through learning and with volitional exploratory movements within the environment. In order to understand more precisely the conditions from which the externalization of percepts arises, Lenay, Canu and Villon (1997) undertook a series of experiments using a highly restricted technological apparatus, a photo sensor fixed to the index finger. By allowing only certain movements, they were able to track the conditions under which externalized spatial perceptions arose. They report that with free movement in three dimensions comes a "spectacular ability to recognize forms…accompanied by an exteriorization of the percepts, which become objects located in space". This kind of experience is an example of technologically mediated direct perception.

What additional design opportunities are opened by these insights into perceptual augmentation? In rehabilitation, the objective is to replace missing or faulty sensory perceptions by means of technological augmentations that map elements of the normally perceivable spectra to alternative sense modalities. In Bach-y-Rita's view, sensory substitution allows for missing sensory information to reach its respective processing area of the brain though transduced by other sensory surfaces. The approach taken here differs in that it rejects the understanding of perception based on sensory modalities (Krueger 2006). It interprets this work not as a substitution of skin for eyes, for example, but as an indication that interfaces can be developed that allow the direct perception of environmental dimensions that are not facilitated by an individual's biology.

In my view, these devices serve as the 'proof-of-concept' for research into 'prosthetic perception'. From the standpoint of the blind or deaf individuals, these devices have allowed for entirely new experiences of the world. They suggest that we may be able to fabricate new perceptual dimensions derived from manufactured sensor technologies if the proper methods of interface can be developed. This research agenda is being pursued at the Rensselaer Polytechnic Institute by the author.

Bach-y-Rita (1972) notes that the brain mechanisms that underlie sensory substitution systems should be similar to, or identical with those that could be used for "sensory augmentation or supplementation" and that to "constitute such a system it is only necessary to present environmental information from an artificial sensor in a form of energy that can be mediated by the receptors at the human-machine interface". He also suggests that it is "possible to provide information from any device that captures and transforms signals from environmental sensors (Bach-y-Rita, *et al* 2003).

An enactive understanding of perception allows – perhaps even asserts – that the reality of our lived experience can be altered, shaped and enriched by technologies (Stewart, Khatchatourov and Lenay, 2004). As Maturana (1997) states, "(Changes in the dimensions of structural coupling) can occur through design, in the intentional use of prosthetic means that create new dimensions of interactions for an organism which thus become new sensory domains for them". The present effort to expand human awareness by design asserts that this new understanding of the process of perception will allow for the design of specific devices that are able to include spectra that are not normally available to human perception. It is suggested that this can be accomplished by means of technological devices that facilitate a structured relation between the output of the interface and volitional movements conditioned on the opportunity to develop these skills over time.

The objective of this research is not in the production of the 'prosthetic senses' as technological artifacts or as isolated sensations. Its success lies in the new percepts augmenting and modifying the apprehended world given by existing perceptions.

Several initial devices have been developed and are undergoing testing in the lab. The first device was produced as a simple demonstration for a seminar conducted by the author. It consisted of a photosensor mounted on a pair of reading glasses with output via a small vibrator mounted to one of the lenses. It was based on the work of Lenay, Canu and Villon (1997) and Borg (2001) as described above. The results were poor because the glasses interface was intolerable for the 10-15 hours of use that are typically required for the externalization of the percept. The failure of this rather simple device is a reminder that human factors such as comfortableness and tolerability can not be overlooked when that objective is the eventual perceptual transparency of the device and its incorporation.

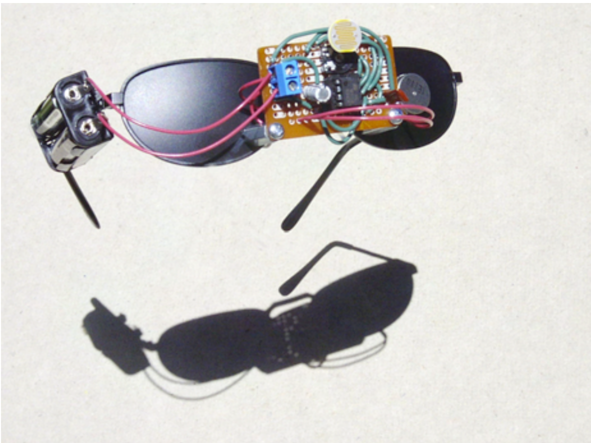


Figure 1. The Glasses that Failed

A device fabricated by the author and his students was based on an eight direction automotive compass and took the form of a belt in which information about the orientation with respect to the earth's magnetic field is conveyed by means of small vibrating motors such as those used as alerts in mobile communications. This device was tested in the field and was found to be, under certain conditions unreliable as an orientation aid. This was due to the prevalence of large magnetic fields surrounding power transmission lines and the fields generated by the large electric motors of suburban trains. The difficulties in using this device were not due to its unsuitability as an instrument for perceiving magnetism, but to its designation as a device for a specific purpose – navigation and orientation. That this decision was made with intellectual knowledge of the prevalence of magnetic fields in industrialized cultures but without an understanding or experience of their ubiquity, strength or dynamics illustrates the thesis that there is much of the world that remains beyond our experiential apprehension. It also illustrates specifically the difference between knowledge gained through representation – scientific visualization – and the embodied knowledge that comes from perceptual experience.

Figure 2. Compass-Belt

Another wearable device was developed to perceive the smaller more intense fields generated by a complex composition of rare earth magnets. It takes the form of a probe built into the body of a pen.

Figure 3. Pen-like magnetic interface

The device was designed to take advantage of culturally conditioned experience with these artifacts and is a direct small scale magnetic implementation of a blind person's cane. Indeed we are 'blind' in relation to magnetic fields. With active exploration, the device allows for a 'focal awareness' of the strength and location of magnetic fields. The quality of the experience is not unlike the awareness that one has in using other probing instruments but is not conveyed by the resistance of physical objects to movement.

An example of a slightly different interface can be seen in the third example. It consists of a glove that contains a fingertip sensor that picks up fluctuating electric fields by induction. These signals are amplified and applied to the back of the finger by a vibro-tactile transducer. While, the vibrations are applied to the skin that on the back of the fingers is relatively insensitive to location, the vibrations are also conducted by the bone that lies immediately below. This makes the vibration difficult to localize precisely. In something akin to a tactile ventriloquist effect, a tingling is felt about the fingertip; sensation is thrown to the point of focal awareness. Like a bone conducting headphone, where the sound is conveyed by bone conduction without interfering with normal hearing, this vibration does not mask the tactile sensations from the fingertip

The glove is intended to provide tactile feedback on the orientation, strength and frequency of the fields which the left index fingertip explores. The laws of sensorimotor contingencies can be developed through repeated use. The experience of these fields will be integrated into the full range of other senses in the context or normal activities. While preliminary results are quite promising the glove is presently undergoing testing and evaluation.

Figure 4. Glove Interface

The prototypes shown in this paper are intended to be the preliminary explorations into the possibility of augmenting human perception with the technological interfaces necessary to expand perception to phenomena that are not now directly available to us. In principle, this approach could be extended to a vast range of sensing technologies. In addition, it makes possible a range of experiments that test and extend our understanding of human perception of the 'world' and of the 'self'. Inspired by the 'sensory substitution' work of Paul Bach-y-Rita and colleagues, it aspires to allow the 'bringing forth' of a phenomenally richer experience by augmenting the sensory flux with new patterns that can be integrated into experience. The work is undertaken primarily as an implementation and testing of enactive perceptual theories, but the projects epistemic dimensions may be augmented by functional and instrumental uses. It is hoped that this may also open new aesthetic territories to investigation.

References

[1] Borg, E., Rönnberg, J., and Neovius, L., "Vibratory coded directional analysis: evaluation of a three-microphone/four-vibrator DSP system", Journal of Rehabilitation Research and Development, Vol. 38 No. 2, March/April, 2001 available (7.25.07) at: www.rehab.research.va.gov/jour/01/38/2/borg382.htm

[2] Bach-y-Rita, P. *Brain mechanisms in sensory substitution*, Academic Press, New York, NY., 1972

[3] Bach-y-Rita, P., Tyler, M., and Kaczmarek, K., "Seeing with the Brain", International Journal of Human Computer Interaction, Vol.15, pp.285-295. 2003

[4] Glanville, R. "Construction and Design", Constructivist Foundations, Vol.1 No.3, pp. 103-110, 2006

[5] Krueger, T. "Nonsense", Technoetic Arts, Vol 4, no.3, pp. 183-192

[6] Lenay, C., Canu, S., Villon, P. "Technology and Perception: The contribution of sensory substitution systems", Second International Conference on Cognitive Technology (CT '97), Aizu, Japan, pp. 44-53, 1997

[7] Maturana, H. (1997), "Metadesign", Essay available (7.25.07) at www.inteco.cl/articulos/metadesign.htm

[8] Maturana, H. and Varela, F. *Autopoiesis and Cognition: the Realization of the Living*. Boston Studies in the Philosophy of Science No. 42, D. Reidel Publishing Co, Dordecht, 1980

[9] O'Regan, K., and Noë, A. "A sensorimotor account of vision and visual consciousness", Behavioral and Brain Sciences, Vol. 24, No. 5, pp. 939 -1031, 2001

[10] Philipona, D., O'Regan, J., Nadal, J. "Is there something out there? Inferring space from sensorimotor dependencies", Neural Computation, Vol.15, No.9, pp. 2029 – 2049, 2003

[11] Stewart, J., Khatchatourov, A., Lenay, C. (2004), "Enaction and Engineering", paper submitted to the on-line Conference on Enaction available (3.30.07) at www.interdisciplines.org/enaction/papers/7/version/original

[12] Stoffregen, T. "On the Nature and Perception of Depictions" 2003

[13] Stoffregen, T., Bardy, B. On specification and the senses. *Behavioral and Brain Sciences*, *24*, 195-261, 2001

[14] Thompson, E. and Varela, F. "Radical Embodiment: neural dynamics and consciousness", Trends in Cognitive Sciences, Vol. 5, No. 10, pp. 418-425, 2001

Interactions of Evolving Artificial Creatures

Nicolas Lassabe Hervé Luga Yves Duthen

IRIT/UT1 UMR 5505
21 Allée de Brienne
31042 Toulouse, France
E-mail: [lassabe,luga,duthen]@irit.fr

Abstract

Cognitive sciences aim at modelling life behaviours. In Artificial life and Artificial Intelligence, there are several approaches to model living behaviours. The approach proposed here aims at illustrating that when the process that regulates the evolution of creatures takes into account the interaction with the environment, the so-obtained artificial creatures presents really sensory-motor behaviours very near to those exhibited by real elementary creatures. In this paper, we designed creatures able to evolve toward motor skills such as walking, jumping, climbing and skating. Such complex motor behaviours are emerging from an adaptive process to the environment taken in charge by evolutionary algorithms.

1 Introduction

In this article we propose a new step in virtual creature evolution. Brooks was one of the first to propose to use artificial life simulation to generate evolving robots and creatures [1]. This article is a resume our previous articles [2, 3]. In recent years, many articles on evolving creatures have been released [4, 5, 6, 7]. They often focus on the morphologies and behaviours of one or two creatures but do not take into account the complexity of their environment. In contrast, some simulations focus on the interactions of population dynamics but deal only with much simpler creatures [8]. Our approach tries to merge both approaches to evolve complete complex creatures in complex environment. After a brief discussion on previous approaches, we will show how evolved creatures are able to cope with various situations caused by a complex environment.

The present work is organised as follows: in the next section will present the works about evolving creatures. In the thrid section of this article we will present a brief description of our creatures. We will particularly emphasize our proposition to use a classifier system as a tool to model complex activation function. We will then present a set of benchmark tests to evaluate our architecture and we will describe the results. At the end we discuss how various fitness functions have an influence on the phylogeny of creatures and on various ideas for future studies. In the results section, we present the best evolving creatures from our experiments. The study of emerged strategies to resolve the various problems is promising and opens new perspectives. We discuss the influence of fitness on the symmetry and modularity of creatures. To conclude, we present some perspectives of the project and the interest that a complex ecosystem environment presents.

2 Related work

There are many existing solutions for representing morphologies and produce emerged behaviours. One of the first works was published by A. Fukunaga, J. Marks and J.T. Ngo [9]. Karl Sims's creatures certainly remain the most evolved [7]. He has achieved evolutions bound by very few constraints and obtained a large diversity of creatures giving birth to crawling and swimming behaviours. Thereafter many works reproduced his results with different approaches for the representation of morphologies and behaviours, such as the L-System creatures of Hornby [4] and Lipson's Golem project [6]. Their goals were to optimize 3D creatures in relation to a preset fitness inside a physics engine and to reproduce them in the real world. Another approach to obtain more complex behaviours is to optimize a set of 2D creatures in a virtual pool by means of natural selection [8]. Each creature has an amount of energy and tries to survive looking for food in the environment. Dealing with simpler creatures, this approach allows to simulate many creatures at one time. Komosinski did the same experiment in a 3D environment [5]. Recently, the EvolGL project [10] proposed to simulate a virtual world with worms. The evolving worms belonged to herbivorous, carnivorous and omnivorous species and developed different strategies of survival.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

These works distinguish two distinct ways to optimize complex behaviour evolving creatures: with a preset fitness or by means of natural selection. In these works, Miconi reproduced the Karl Sims's creatures and improved the co-evolution [11]. One promising perspective for the morphology of creatures could be artificial embryogeny to have a more complex representation of creatures [12]. However, much could be improved to offer better results in more complex environment [13, 14, 15]. The works based on natural selection are becoming achievable and will certainly be interesting to have more adaptive solutions [10].

3 Environment and architecture

3.1 Morphology and Behaviour

In present work, our creatures use a morphology based on *graphtals* similar to those from Karl Sims's ones [7]. The genotype represents the creature's development. For each creature we define a root node. Each node contains the information of one block. Blocks are represented by solid 3D cubes. The modularity is very visible in our results. We choose not to explicitly represent the symmetry in our creatures as in Karl Sims to study how the evolution could make it appear. The behaviours of the creature emerge from the classifier system[2].

3.2 Genetic operator and Genetic Algorithm

The mutations are operated on the graph, hence if a mutation is operated on a node recursively reproduced, this mutation will affect all the blocks issued from this node. This allows our structure to exhibit modular properties such a pattern repetition. For the crossover, we use the *graphts* method as in Karl Sims's works [7] to reproduce two creatures. For the genetic algorithm, the population of the simulation is initialized to one hundred individuals. Each creatures genotype is randomly. Each individal's performance is then evaluated in the simulated environment during 40000 steps (40s).

4 Results

We here show the ability of our system to evolve creatures in different situations. In general, in our experiments, the fitness tries to maximize the distance traveled by a creature. All the creatures are recorded on videos[1].

[1]A video with the best creatures is on the first authors's home page: http://www.irit.fr/∼Nicolas.Lassabe/videos/alife2007.mov

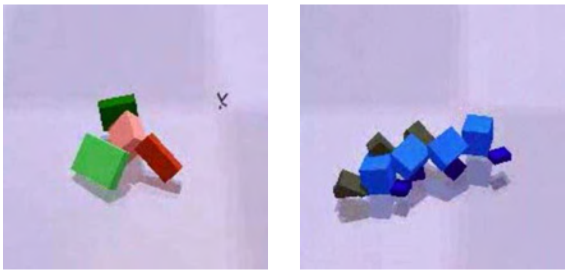


Figure 1: Evolving creature walking like a tripod and evolving creature walking like a crabe.

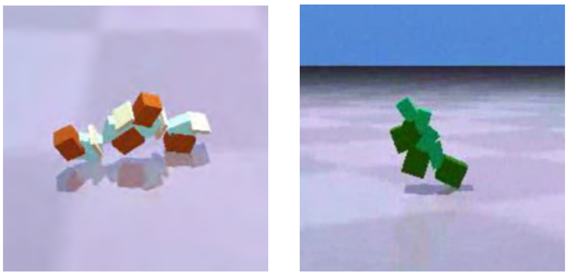


Figure 2: Look like a robot and a creature with a human profil

4.1 Walking on a flat floor

The first experiment is the traditional walk (Fig. 1). This original experiment is similar to the work of Karl Sims [7], who presented the first one. Many works try to reproduce it. Its goal is to optimize both the morphology and the behaviour of a creature which has to walk on a flat floor environment. In this experiment, our fitness F is not the distance traveled d but the average speed during the time t of evaluation of the creature. $F = d/t$ One of the best results that we get is a tripod. This creature is not symmetric: it has only three blocks for legs and one for the body and is not modular (no repeated block). Various other creatures used interesting strategies to move like a crab or have kinds of legs or rolling.

4.2 Walking in a specific direction

In the experiment, we keep the same simulation and parameters but we change the fitness function. In this case, we indicate a specific direction for the creature, for example, to follow the x axis. The new fitness is $F = |x|/t$ One might imagine the results would be the same. In fact, the creatures are optimized to go in this direction and this is manifested by a perfect symmetry of the best creatures. An interesting feature our creatres's morphology of creatures is that they are composed of a modular body. This modularity gives them a great efficiency. Among the best creatures are kinds of worms with some legs allows than to walk with long steps.

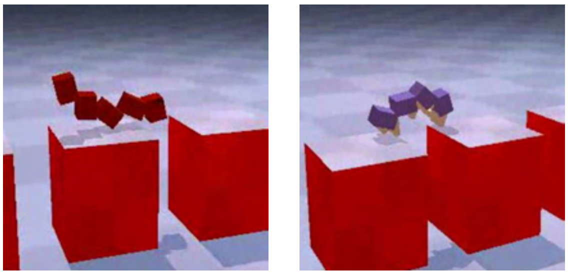


Figure 3: Two evolving creatures on cubes.

4.3 Thought trenches

In these experiments, we isolate the creature on a small cube (Fig. 3). Some other cubes are aligned with the first one separated by trenches. The creature's goal is to join the last cube. Many strategies are possible like jumping over trenches. Or, if the creature is tall enough; it could crawl to the next cube. The interest of this experiment is that this situation often appears in reality and is a difficult task to resolve. It is easy to improve the difficulty of this task changing the width of the trenches. It is possible to alternate the positions and heights of cubes. The best creature uses its legs to cross the holes.

4.4 Stair Climbing

The environment of the creature is often a flat floor. To increase the difficulty, we try to see if an evolution could generate a creature able to climb a stair. The difficulty could be increased by the height of steps. In the first test, we set up the height of the steps to half the average size of creatures and the width as the average width of the creature. The fitness is defined to indicate the direction to the top of the stair and optimizes by speed to which the creatures go. Stairs are everwhere in daily life, and can be of sereval diffirent types. It is a difficulty for a robot. Using evolving creatures to see the different strategies to climb a stair is an interesting approch. The morphology, the balance and the synchronisation of movements are difficult to design to work on any stair. The best creature can climb to the top of a stair quickly. The strategies employed are different. Some creatures which move like a worm are able to climb almost two steps at a time. Others are perfectly designed to jump step-by-step. Once again, the bodies of the best creatures are modular . Morever the creatures are also symmetric. If a creature has a good potential to climb but is not totally symmetric, the evolution will optimize the creature in this way. To improve the difficulty of the experminent, we increase the height of the step and run a new evolution. The best creature uses a new strategy to climb the step using flat legs.

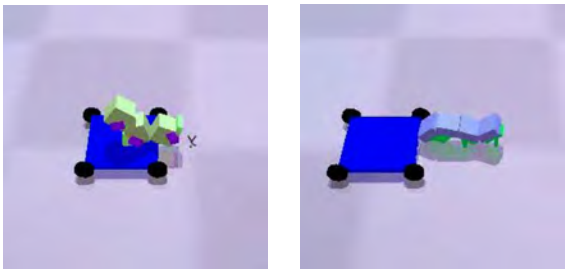


Figure 4: Evolving creature skating by jumping on the skate and evolving creature pushing the skate.

4.5 Skating

The interaction within the environment is an important issue in improving the complexity of evolving creatures. In general, we underestimate the capacity and the power of adaptivity of evolving creatures. Many experiments need to be tried to find out their real limit and their potentialities. In this experiment, we define a kind of skateboard composed of four wheels and one large board. The goal is to see creatures could use it whether to move in some direction. For this, the creatures fall down close to the skateboard. The creature and the skateboard are two different objects and do not have any connection between them. The function fitness is defined to maximize the distance covered with skate ds. $F = d + ds - (d - ds)$ Many strategies shows employed: some creatures jump on the skate to slide with it, others push it and run behind it and others go up and push from the back to move. These results show the ability of the evolving creatures to be adapting to the environement (Fig. 4).

5 Discussion

These results prove the possibility to produce emerged behaviour from classifier systems for evolving creatures in various activities. The next step is to emerge our system inside our virtual ecosystem with plants and various lifeforms. We have already started this implementation [3].

5.1 Why ecosystem?

The realization of an ecosystem can allow the generation of lifeforms able to adapt to a dynamic environment, the specification of species and the emergence of life cycles. The interest of simulating such virtual ecosystems are several. In the field of biology, it could lead to a better understanding of the evolution and extinctions of species as well as of some mechanisms of life [16]. The evolution of morphology and behaviour in relation to a specific environment has an interest for the design of robots and their behaviours [17, 6, 18]. A complex environment should generate adapted survival strategies during evolutions cycles.

Interesting questions are whether different evolution scenarios can emerge from an identical initial environment and whether the environment complexity is a relevant criterion on the emergence of different survival strategies [13, 14, 15]. If that were the case, how does one evaluate complexity? What are the elements necessary for the emergence of complex life forms and behaviours? Is the use of a complex physics one of these elements or can biological complexity be supported by simpler physical models? What is the importance of the interactions? These are open questions, which may be answered through experimentation with a complex artificial ecosystems.

6 Conclusion and Perspectives

Following the work from Sims and from Brooks, we presented here creatures that are able to adapt to their environment and develop motor skills by themselves. Such creatures exhibit emergent behaviours in the sense that the behaviours are not explicitly represented by explicit rules but emerge globally from a small set of internal rules including adaptation. The next step for our project will be to propose some paths to improve the interaction within the environment, and also to confront the evolved creatures with more complex situations to exhibit their adaptation abilities. Introducing an evolving language between the creatures could be interesting and would allow a significant increase in the number of evolving strategies. It could also be very valuable to change our morphogenesis definition allowing genotypic re-use. This could lead us to an evolving system based on the artificial embryology paradigm where a compact genetic code could lead to the growth of complex structures.

References

[1] R. Brooks, "Artificial life and real robots," in *European Conference on Artificial Life*, 1992, pp. 3–10.

[2] N. Lassabe, H. Luga, and Y. Duthen, "A New Step for Evolving Creatures," in *Artificial Life - ALife'07*. IEEE, 2007, pp. 243–251.

[3] ——, "Evolving creatures in virtual ecosystems," in *16th International Conference on Artificial Reality and Telexistenc e*, 2006.

[4] G. Hornby and J. B. Pollack, "Evolving l-systems to generate virtual creatures." *Computers & Graphics*, vol. 25, no. 6, pp. 1041–1048, 2001.

[5] M. Komosinski, "The world of framsticks: Simulation, evolution, interaction," in *VW '00: Proceedings of the Second International Conference on Virtual Worlds*, 2000, pp. 214–224.

[6] J. B. Pollack, G. S. Hornby, H. Lipson, and P. Funes, "Computer creativity in the automatic design of robots," *Nature*, vol. 406, 2000.

[7] K. Sims, "Evolving virtual creatures," in *Siggraph'94*, 1994, pp. pp. 15–22.

[8] J. Ventrella, "Designing emergence in animated artificial life worlds," in *VW '98: Proceedings of the First International Conference on Virtual Worlds*, 1998, pp. 143–155.

[9] A. Fukunaga, J. Marks, , and J. Ngo, "Automatic control of physically realistic animated figures using evolutionary programming," in *Proc. of Third Annual Conference on Evolutionary Programming*, 1994, pp. 76–83.

[10] S. Garcia Carbajal, M. B. Moran, and F. G. Martinez, "EvolGL: life in a pond," in *Artificial Life XI Ninth International Conference on the Simulation and Synthesis of Living Systems*, 2004, pp. 75–80.

[11] T. Miconi and A. Channon, "The n-strikes-out algorithm: A steady-state algorithm for coevolution," in *the IEEE Congress on Evolutionary Computation*, 2006.

[12] K. O. Stanley and R. Miikkulainen, "A taxonomy for artificial embryogeny," *Artif. Life*, vol. 9, no. 2, pp. 93–130, 2003.

[13] C. Adami, C. Ofria, and T. C. Collier, "Evolution of biological complexity," *PROC.NAT.ACAD.SCI*, vol. 97, p. 4463, 2000.

[14] A. K. Seth and G. M. Edelman, "Environment and behavior influence the complexity of evolved neural networks," *Adaptive Behavior*, vol. 12, no. 1, pp. 5–20, 2004.

[15] A. K. Seth, "The evolution of complexity and the value of variability," in *ALIFE: Proceedings of the sixth international conference on Artificial life*, 1998, pp. 209–218.

[16] G. Abramson, "Ecological model of extinctions," p. 785, 1997.

[17] D. Floreano and F. Mondada, "Automatic creation of an autonomous agent: Genetic evolution of a neural-network driven robot," in *Proceedings of the Conference on Simulation of Adaptive Behavior*, 1994.

[18] S. Nolfi and D. Floreano, "Synthesis of autonomous robots through evolution," 2002.

Embodied Listening Performances Reveal Relationships Between Movements of Player and Listeners

Marc Leman Frank Desmet Frederik Styns Leon van Noorden Dirk Moelants
Dept. Of Musicology, Ghent University
Blandijnberg 2, B-9000 Ghent, Belgium
E-mail: Marc.Leman@UGent.be

Abstract

The movements of a guqin player and several listeners were recorded using kinematic sensors. Movement velocities were extracted to test whether listeners correlate with each other and with the player. The experiment revealed that listeners tend to mimic action events that underlay music. The findings provide evidence for the hypothesis that music perception has roots in action. The study offers a new methodology for studying the action-based component of music perception in several domains of music research.

1. Introduction

In embodied listening experiments, the subject is encouraged to get corporeally involved with a musical stimulus. There is evidence that music signification practices can be understood from the viewpoint of action, and that perception of music spontaneously involves action [1]. This evidence can be linked with recent theories that assume shared neuronal codes as the underlying structure for the coupling of perception and action [2, 3]. However, many questions remain unanswered, for example: how listeners signify music through action, how they engage with music, and how they engage with each other in terms of corporeal movements.

In the present study, the focus is on the measurement of corporeal movements of a single musician (called player) and a group of listeners. The aim of the paper is to explore the hypothesis in embodied music communication that movements between player and subjects, and among subjects can become synchronized. The methodology has a focus on movement patterns and comparison of movement patterns among different subjects and different listening sessions.

In the present paper, we use movement infra red camera capturing systems and joy-stick sensor technology to monitor (a) how a musician moves while playing and (b) how listeners move while attuning or harmonizing with the perceived music. From the measurement of these movements, we extract information about the origin and the nature of the emulated behavior using standard multivariate statistical techniques. Velocity is a core cue of movement, for that reason we focused on the extraction of movement velocity. This focus is in line with several other studies in which movement velocity has been used as a cue.

Figure 1. Player and motion capturing

2. Experimental Setup

In order to identify the movements that underlay the actions of player and listener, audio, video and kinematic data of three short (<30 sec) pieces (P1,P2,P3) of Chinese guqin music were recorded. Two of these pieces, P1 and P2, have a rather fluent melodic line which is clearly structured. In contrast, P3 has a more rubato and narrative character with a less fluent melodic line. The kinematics of the player was recorded with an infra red camera system (Figure 1). Likewise, thirty subjects, unfamiliar with this music, and with Chinese music in general, were asked to

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

move a joystick while listening to paired repetitions of the audio recordings (Figure 2).

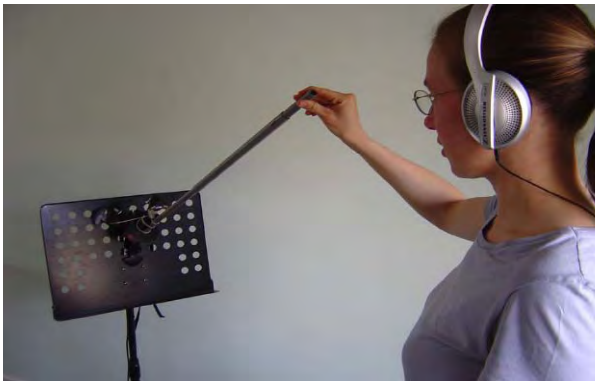


Figure 2. Listener responding to music via joystick

This task was done in four sessions that immediately followed each other and induced learning. In the first session, subjects were asked to move the joystick while listening to the musical audio stimulus. Before the start of the second session, subjects were invited to listen to the fragments in random order and as many times as they wanted. When finished, they were asked to move to the same sequence of musical fragments as in part one. Before the start of the third session, subjects were asked to listen to the musical fragments, now using a visual representation (pitch versus time graphs) of each musical fragment. Again the subjects were allowed to listen/look in random order as many times as they wanted. When finished, they were asked to move as in session one. Before the start of the fourth session, subjects were asked to watch/listen a video of the player's performance of the three pieces. The subjects could watch/listen in random order and as many times as they wanted. When finished, they were asked to move as in session one.

The listener's kinematics were recorded and after each session, the listeners had to fill out a questionnaire that probed the subject's self-assessment of the performance. The main question was “Do you have the feeling that your movements were in accordance with the music?” (‘very badly’ to ‘very well’, asked in session 1 to 4, for each piece). All subjects were thoroughly verbally instructed about what they were expected to do at the beginning of the experiment. The subjects were also given a short written instruction that repeated what the instructor said.

3. Data Analysis

Movement velocity is calculated as the displacement, in three dimensions (player) or in two dimensions (listener), during non-overlapping time windows of 250 ms. Movement velocities are used in two types of measurement.

First, intra-subject comparisons address the subject's ability to repeat the movement velocities during subsequent repeated listening of a piece (e.g. P1P1) during a session. A correlation between both movement velocity patterns at the $p < 0.01$ significance level was used as a criterion for making the distinction between successfully and non-successfully repeated movements. This choice was based on a detailed inspection of a number of correlated movement velocities.

Second, inter-subject comparisons access the relationship between the movement velocities of different listeners. Similarly, the computational comparison of the movement velocity vectors is based on correlation. In addition, for inter-subject comparisons, a dynamic time warping technique is used. In the latter, the listener's movement velocities are first warped onto a reference pattern, namely, the player's movement velocities. Correlation analysis is then based on warped patterns. This dynamic time warping (DTW) accounts for the fact that movements of the subjects can be somewhat shifted as compared to the player 's movements (anticipated or lagged).

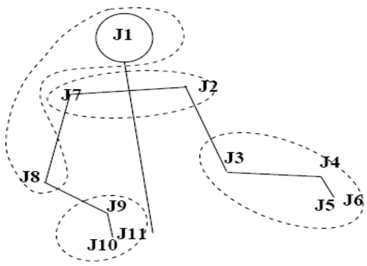
Moving along with music may be *in tune* or *out of tune*. In tune would mean that the listener has the feeling that the movements somehow are in accord with the music. Out of tune would mean that they are not in accord. The latter would occur, for example, when no movements were made, when hesitating movements were made, or when the listener had the feeling that the movements were too late or too early, and so on. In order to be able to address the action component in relation to music, it was important to try to access this distinction in an objective way. To address this issue in a quantitative way, it was assumed that listeners who found a proper motor strategy to express their movement *in tune* with music (during 30 seconds) would not radically change their strategy in a second performance that immediately followed the first one, while listeners, who found that their motor articulations were *out of tune* with music would be inclined to change strategy and therefore have a different performance in the second trial. Thus, high correlation between successive performances (e.g. P1P1) would indicate *reliable performance*, while low correlation would indicate *non-reliable performance*. The measurement of the correlation between the two repetitive performances thus offers an objective criterion for analysis.

4. Results

In the course of this study, there were no effects of gender, age, musical background, listening hours per week, and absolute pitch perception. Therefore, these effects can be discarded from further analysis.

Dynamic time warping (DTW) was used to compare movement velocities of the listeners with movement velocities of all joints of the player. All reliable performances were selected for this DTW analysis. The warp path was restricted to a maximum of 750 ms from the diagonal (which corresponds with 3 time units). The path with the best resulting correlation and the lowest cost was chosen. The resulting cost/correlation analysis shows that there is a relation between correlation and cost. The higher the DTW correlation obtained, the lower the cost it takes to achieve this. This analysis shows that the movement of the right shoulder of the player has the highest average correlation with the lowest cost (Figure 3b). There is also an indication that the player's movements can be classified into 4 different groups: right hand; left elbow and left hand; head and right elbow; shoulders (Figure 3a).

(a)



(b)

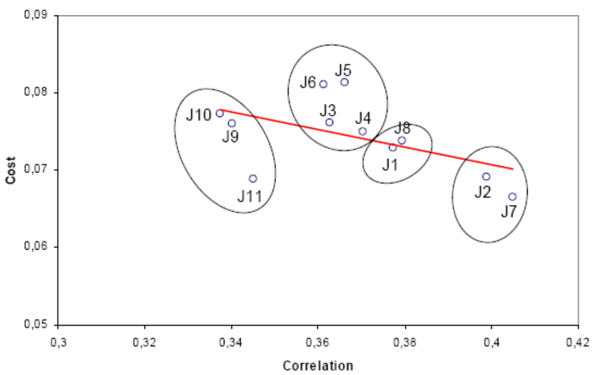


Figure 3. (a) Grouping of joints of the player, (b) Correlation and cost between different joints of the paper and the response of the listeners

Based on the above results, the movement velocities of the listeners were then time-warped with one single movement marker, the right shoulder. General Linear Modelling (GLM) for Repeated Measures was then used to model the number of significant listener-player correlations for *reliable* and *non-reliable* performances over the four sessions. For the reliable performances, there is a significant effect over session ($p = 0.005$) and no significant interaction effect between session and piece. The latter means that the correlation of the movements with those of the player increase over session and that the pieces follow different trends. In contrast, for the non-reliable performances there is no significant effect over session. This result for non-reliable movements confirms the separation between the two types of movements (reliable versus non-reliable).

When listeners move in tune with the music, one may expect that inter-subject correlations can be observed. In the present study, the number of inter-subject correlations, based on all performances (reliable as well as non-reliable) was low. GLM for Repeated Measures reveals a significant difference between pieces. This difference can be attributed to the difference between P1, P2 on the one hand and P3 on the other hand. The analysis shows a significant effect over session ($p < 0.001$) and a significant interaction between session and piece ($p < 0.001$). Data analysis based on warped movements shows inter-subject correlations in about 50% of the movements for P1 and P2, and about 30% for P3. The rubato narrative character of P3 seems to be less predictable in terms of action events, and the movements seem to have a less pronounced inter-subjective component.

Concerning intra-subject correlations, the number of reliable performances increases over sessions. GLM for Repeated Measures reveals a significant difference between pieces ($p < 0.05$). There is a significant effect of session for a linear model ($p = 0.005$) and there is no significant interaction between session and piece. Learning works best for P2 and P1, that is, from above 40% in S1 to about 70% in S4. Instead, for P3, there is a smaller increase. This means that, for reliable performances, subjects tend to improve in their individual learning path for all pieces. Seen in the context of the inter subject results, it can be concluded that this learning effect is based on individual interpretations rather than on common inter-subjective interpretations. In other words, over sessions, the individual interpretations improve, but they drive away from a common inter-subjective interpretation

5. Discussion

A direct comparison of the movement velocities of the player with the movement velocities of listeners revealed some interesting relationships.

First, it was observed that the movement velocities of the listeners' arm movements tend to correlate with the movement velocity of the player's shoulders. This finding seem to imply (i) that listeners can embody the musical stimulus by translating the perceived sounds into actions, and (ii) that action components of listeners are related to action components of the performer. It should be noted here that guqin music is particularly suited for studying these relationships because the technological mediation between the player's bio-mechanical energy and the sound energy is very straightforward. Due to the fact that the instrument has neither frets nor bow involved, the sonic trace in music can be considered a direct image of the player's action. Therefore, it could be argued that the sonic traces of the guqin facilitate the transformation into action patterns that relates to the player's action pattern. However, listeners do not emulate pitch modulation movements because if they did, there would be a correlation with the movements of the player's finger. Instead, it was observed that listeners tend to emulate the player's shoulders. The role of the shoulders is particularly striking and it may point to the fact that listeners emulate the player's musical *intention* rather than the player's technical playing gestures. The decoding can only be explained by the fact that listeners have a capacity to emulate the action patterns that are contained in music. The findings suggest that through action, the listener builds up a corporeal understanding of the musical patterns. Music, under certain circumstances, would thereby operate as a channel through which action patterns can be transmitted.

Second, it was observed that the music-driven movement velocities of the listeners tend to correlate with each other. Beat and rhythm are often concerned as the main sources for musical embodiment; however, the musical stimuli used in the present study did not contain a very clear beat. Despite these conditions, a growth of correlation among the movement velocities of the listeners could be observed over sessions. Therefore, it can be concluded that listener-listener correlations may be based on the decoding of expressive features in music, rather than on mere beat following. The present study suggests that embodied listening is prone to learning and that in a condition where learning is stimulated, the quality of the listener's movements tend to improve. This learning, apparently, may be strongly depending on the musical character, as the difference between P1-P2 and P3 seem to indicate.

Finally, it was noticed that Piece 3 revealed an interesting difference between intra and inter subject analysis. The intra subject analysis suggests that listeners tend to improve for P3 over session, while the inter subject analysis suggests that listeners tend to drive away from a common interpretation. The two trends can be explained by the fact that listeners develop their own embodied listening solution to P3 in a consistent way but this listening solution differs among listeners. The lack of a shared inter subjective embodied listening solution may be due to the narrative character of the piece.

To sum up, the present study suggests that embodied listening is based on a music communication model which consists of (i) the sonic encoding of actions by the player, (ii) the transmission of these sonic cues (through music), and (iii) the decoding (or corporeal understanding) in terms of action events by listeners who perceive these sonic cues. The transmission of a corporeal code forms part of a mechanism where listeners, in order to bridge the semantic gap between sonic cues and meaning, attribute ‘plausible’ or ‘virtual’ action events to music, which they take from their own action repertoire. In this way, through embodied attuning, music perception is grounded in the listener's subjective experience. This foundation in action will give it a proper inter-subjective aspect, as shown in this study.

6. Conclusions

The present study provides evidence for the theory that music perception is rooted in action. Based on direct measurement of movements, it was shown that listeners tend to mimic action events that underlay music. Embodied listening can be explained by the existence of shared codes for action and perception, allowing the translation from sonic energy via bio-mechanical energy into the perception of action, or action intentions. Clearly, the decoding of action in music draws on a subtle mechanism in which the listener tends to disambiguate the stimulus by body movement. The present study suggests that this mechanism is not necessarily based on knowledge of how to play the instrument, nor is it merely constrained by the bio-mechanics of the human body. It was observed that the (reliable) movements of the listeners' arms correlate with the movements of the player's shoulder. This seems to indicate that there can be a genuine transfer of action patterns from the player to the listeners. “If this hypothesis is true, then it points to the existence of a corporeal code which can be encoded and decoded and which music is able to transmit.

References

[1] Leman, M., Embodied music cognition and mediation technology, Cambridge, The MIT-Press, in press

[2] Hommel, B., Musseler, J., Aschersleben, G., and Prinz, W., “The Theory of Event Coding (TEC): A framework for perception and action planning”, Behavioral and Brain Sciences, 24(5), pp. 849-939, 2001.

[3] Sebanz, N., Bekkering, H., and Knoblich, G., “Joint action: bodies and minds moving together”, Trends in Cognitive Sciences, 10(2), pp. 70-76, 2006.

Externalism and Enaction

Charles Lenay[*] Pierre Steiner[*]
[(*)]*COSTECH - UTC*

E-mail: charles.lenay@utc.fr, pierre.steiner@utc.fr

Abstract

In order to illustrate the relation between the enactive theory of perception and the active externalist view of perceptual consciousness, we present experiments based on minimalist Enactive Interfaces, inspired by the work of Bach-y-Rita. We particularly aim at showing (i) how technological devices can contribute to the enaction *of perceptual consciousness; and (ii) how the space of perception is actively constituted before any spatial inner/outer distinction concerning particular objects and processes.*

1. Introduction

How and when is the distinction between an external world and an inner world constituted? How and when is there a constitution of the space *within which* an object, and the point of view on that object, are both localized?

Answering to these questions is important in order to: (1) understand the functioning of enactive interfaces which lead to the constitution of new perceptual spaces; and, (2) to understand cognitive technologies as technical devices and external inscriptions which participate to our cognitive activity.

The classical representational approaches to perception are *internalist*. They consider perception as the activation or the construction of internal representations that are supposed to be the more or less faithful reflection of an external state of affairs. By contrast, we wish to show that an enactive approach to the definition of interfaces necessarily engages an externalist conception of cognition. This strong form of externalism can be called « vehicle-externalism » [7] or « active externalism » [1; 2; 8; 14], according to which cognitive activity, and in particular perceptual activity, is not especially localized in the brain but equally in the body, the environment and the dynamics of concrete actions. We wish to show here that this strong externalism enables a better understanding of how tools and enactive interfaces contribute to our ways of perceiving and thinking.

First of all, we may remark that the opposition between "externalism" and "internalism" presupposes a space within which the question itself makes sense: a space where objects and living organisms perceiving them are both situated; a pre-given space which makes possible to ask, *afterwards,* if the perceptions are internal states of the organism (in the brain), or if they are also situated in the surrounding world. This opposition between externalism and internalism takes a radical turn with the question of the constitution of a spatial perception. In the following, we will examine this question by means of device, which by virtue of their extreme minimalism induce an exhibition of the active, concrete constitution of a perceptual space.

2. Active perception and supplementation

Our simple and minimalist experimental situations concern two tasks: the *localization* of objects and the *recognition* of shapes by blind and blindfolded subjects, with the help of tactile devices that mediate and create new modes of sensorimotor coupling and thus new perceptual modalities [11].

2.1. Spatial localization

Inspired by the sensory substitution devices of Bach y Rita, our minimalist system is composed of a single photo-electric cell triggering a binary tactile stimulator. When the incident luminosity, within a cone (of about 20°) is greater than a threshold, the tactile stimulus is triggered. Thus, the blindfolded subject receives only minimal information of 1 bit corresponding to the presence or absence of the tactile stimulus.

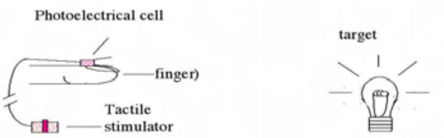


Figure 1. Minimalist experimental device for spatial localisation

In [6], we showed that, even with such a simple device, the spatial location of luminous targets is still possible. The subject initially perceives only a succession of tactile stimuli that accompany his movements. But, along as he becomes familiar with the device, his sensations are progressively replaced by the perception of a target at a certain distance in front of him. The sensory information being only a temporal sequence of binary 1 and 0 with no spatial information, we can say here that perception cannot be grounded merely by a simple internal analysis of it.

Consequently, in such absence of spatial information, the perception of the localization of the target in direction and depth is only accessible by means of *active exploration*. We may say that he minimalism of the device *forces a spatial and temporal deployment of the perceptual activity*. The latter can then be studied on the basis of observable movements.

In such conditions of minimal coupling, it is easy and useful to replace the physical reality which triggers the tactile stimuli according to the orientation of the photo-electric cell, by a motion capture device placed on the finger which defines the position of a receptor field in a virtual space. In such minimalist perceptual supplementation situation, the transformation from the real world to a virtual world makes practically no difference to the subject [23]. The advantage is that in a virtual environment, the perceptual trajectories can easily be recorded, and the conditions of coupling modified by changing, for example, the shape, the arrangement or the number of receptor fields.

In order to maintain his perception of a target placed in front of him, the subject must act continually by displacing the motion-capture device. As soon as the movements stop, the perception disappears. This can easily be understood when we adopt the point of view of the subject. When he is immobile, there are only two possibilities: either he receives a continuous stimulus, or he does not. If he is pointing away from the target, he has only the memory of a perception which fades away. If he is pointing at the target, he receives a continuous stimulation, but this does not lead to the perception of an external object. Spatial perception requires the *synthesis* of a temporal succession of actions and sensations. The spatial exteriority of the target can only be constituted by the possibility of freely and reversibly coming and going around it, alternately leaving and finding contact.

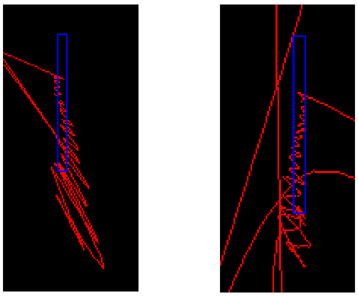


Figure 2. Projection of the pointing movements on 2 planes that include a luminous target.

Subjects perform regular oscillations around the target : small movements of the hand, combined with progressive changes in the position of the wrist. Something is happening since the subjects seek to identify the functional relationship between these actions and the patterns of sensory returns that they produce. The target is localized in direction and depth when the law governing pointing towards it is mastered. This is an illustration of what O'Regan called a "law of sensori-motor contingency" [15]. Any given position of the target corresponds to a particular *sensori-motor invariant*, i.e. a *law* relating sensory feedbacks to the actions performed; this law is stable over and above the evolving actions and sensations.

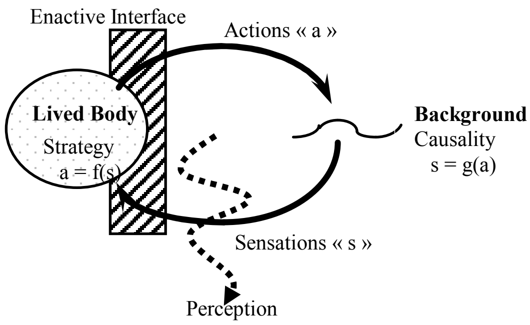


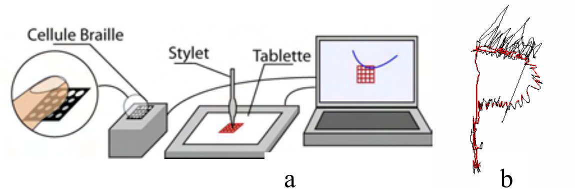
Figure 3 : Scheme of enactive perception (sensori-motor coupling). Via the environment, the actions "a" give rise to sensory feedback "s" : s = g(a) ; concomitantly, the organism implements a strategy for generating its actions and modulating them as a function of its sensations : a = f(s).

When the subject is immobile, and can only perform rotations at arm's length, he can maintain a sense of the orientation of the target, but it is no longer possible to localize the distance of the target; *depth* disappears from the perceived world [9]. Perception depends not only on sensory input, but also on the capacity of the lived body for *action*. In order to trigger a perception, a prosthetic device must be an instrument of coupling which modifies the body by defining new repertoires of action and sensation. Then, perception is more the result of dynamic coupling between the organism and its environment [24] than an internal representation. This is why we situate perception at the core of the coupling, and not unilaterally within the organism. In this approach, an important distinction must be done between "sensation" and "perception. "Sensation" is defined by the sensory input received by the organism, while "perception" relates to the relation between the sensory feedbacks and a full range of actions.

The perceptual actions of a subject correspond to movements of his "point of view", i.e. the locus from which the object is perceived. The spatial localization of the target is simultaneously a localization of the point of view and of its relative movements to the target. The subject, as an organism in movement, belongs to the space in which it is situated with respect to the target. Perception is an embodied activity; it is the body which makes possible to get a hold on the world. The perception of depth is not to be situated in a purely subjective abstract represented space, but rather in the concrete dynamics of the coupling of the organism to its environment. This same idea can also be observed in the case of task of shape recognition.

2.2. Shape recognition

The "Tactos" system [5] has been developed in order to allow blind persons to have access to digital forms present on a computer screen. It consists in a device for controlling tactile stimulators as a function of the movements of a cursor on a computer screen. It is composed of a stylus, a graphic tablet, a Tactos software, and a matrix of Braille cells of which the movements of the pins are controlled electronically.



a b

Figure 4. (a) The Tactos system. (b). Perceptive trajectory for the recognition of a P.

The cursor of the tablet here corresponds to a small 4 x 4 matrix of 16 receptor fields. When one of the receptor fields encounters one black pixel, the binary activation of the corresponding pin on the Braille cell is triggered. The subject is blindfolded, and moves the cursor by means of an effector (mouse, stylus …). The tactile stimulation is delivered to the *other* free hand, but this does not hamper the perception of the forms.

Even in this minimal version, we observe that subjects are able to perceive forms. These forms are not given to the sensory system as a complete two-dimensional pattern applied to the skin. When there is only a single receptor field, and thus a single sensation at each instant, there is no longer intrinsic spatiality at the level of the input signal. If the subjects succeed in recognizing shapes in space – and they do – this can only be by virtue of an active exploration in the course of which they integrate their movements and the corresponding sensory feedbacks over time. Thus, here again, by limiting the sensory input to just a single bit of information at each instant, we oblige the subjects to *deploy their perceptual activity in space and time*. Virtual reality situation allows to record and to analyse this activity, underlying what we called "perceptual trajectories"(figures 4b).[10; 21].

Perceptual trajectory is both a *recognition* and a *constitution* of the shape. The categorization of the perceptual data as belonging to a shape is achieved by a gesture. This gesture is a *scheme* for the construction of the shape; by such scheme, the categories of understanding apply to the sense-data of intuition.[17] Here, the scheme of assimilation corresponds to a concrete activity, deployed in the space of the displacements of the subject. The schema is realized by a "gestural strategy" which, by using exteroceptive and proprioceptive sensory feedbacks, produces a set of movements that make possible to inscribe the shape and to grasp it as a whole in a unitary gesture of anticipation.

By studying the dynamics of perceptual trajectories, we observe the concrete activity of the constitution of a shape in perception. "Imagination", defined as the capacity to produce images, is thus situated at the core of the perceptual process itself, as the capacity to produce the shape necessary for its perception. It is clearly an illustration of the understanding of perception as enaction – i.e. *production* of the perceived object – and *not* as re-presentation.

We may also note that the proprioceptive perception and memory of absolute position is imprecise for the subject for plotting the positions of the hand that holds the effector (mouse, stylus…) in egocentric X-Y co-ordinates. It is thus quite impossible for the subject to scan the whole field of the screen, and to integrate the points of stimulation in order to construct a mental image of the shape. When he inadvertently leaves the contour of the shape, he is immediately "lost", and cannot even proprioceptively return to the last point of contact with the form. He starts out with large-scale exploratory movements, but as soon he obtains a contact with a line, he converges to a *micro-sweeping* movement of small amplitude around the source of stimulation. This process is truly an operation of localization: the position of a fixed spatial singularity is *constituted* by a stable anticipation of the tactile stimulus according to the movements of the receptor field. At the same time, the micro-sweeping movement enables the subject to identify his own position, not in absolute co-ordinates but relative to the form that he is exploring and perceiving.

This localization of the point of action is the condition for being able to perceive oneself in the space created by the technical mediation.

3. Externalism

In an internalist approach of spatial perception, one would have to imagine that the localization or the recognition of shapes is the result of an internal calculation distributed over areas of the brain. Thus, the space of perception is distinct from the physical space of concrete actions. By contrast, in the experiments we described, there is only a single space for the perceptions and the actions. Actions, moreover, taking place in the same space as the perceived objects. These very simple experimental situations provide both a precise definition, and a concrete functioning example, of the explanatory schema of perception as enaction. Each of the two technical devices we presented brings forth a novel perceptual modality. The space of this perception, which is specific in each case, is actively constituted according to the possibilities of actions at the disposal of the subject. For the subject there is, each time, conjoint coming-forth of a perceived object and a point of view on that object.

The perception of the object happens in the same space where the object is – an *enacted* space. It is from this enacted space that the distinctions between the inner and the outer can then make sense, but in a new way: is *inner* everything which moves with my point of view, whereas is *outer* what my point of view moves in comparison with. Space appears as the systemic structure of reversible displacements of this point of view, and of objects relative to the point of view.

Thus, with respect to the ontological question asked at the beginning of this article, the concept of spatial perception defended here is, in a way, neither externalist nor internalist, the space of perception and its contents being constituted in the coupling between a living organism and its environment. It is only on the basis of an "*inbetween*", the relation between both, that a perceptual space is emerging, i.e. a lived world for the organism. Space is the form of this coupling, the structured domain of the invariants that can be constituted. However, if one situates oneself in this space in order to distinguish and localize the point of view with respect to the objects of perception, our approach clearly becomes externalist in the sense that the perception of objects does not occur *behind* the point of view, but *in front of it,* in the very same space where the point of view moves itself.

4. Conclusion

This externalism would be an appropriate epistemological and theoretical framework in order to account for the effectiveness of situated cognition. Technical devices and environments transform our possibilities of action, and thereby transform our lived experience, offering new capacities for perception, imagination, memory and reasoning. Similarly, an externalist framework allows a fruitful dialogue between phenomenology and psycho-physiology. The conventional approach consists of *presupposing* a separation between internal lived experience and external objectivity, which leads only to the search for neuronal correlates of consciousness. By contrast, for externalism, this separation is actually constituted during the very process of concrete activity. As a result, the space of lived experience is co-extensive with the space of action and perception.

References

[1] Block, N., "Review of Alva Noë's Action in Perception", Journal of Philosophy, CII, 5, pp. 259-272, 2005.

[2] Clark, A. Being there: Putting brain, body and world together again. Cambridge, Ma.: MIT Press, 1997.

[3] Davidson, D. "What is present to the Mind", in Subjective, Intersubjective, Objective. Oxford University Press, 2001.

[4] Dretske, F. Naturalizing the Mind, Cambridge: MIT Press, 1995.

[5] Hanneton, S., Gapenne, O., Genouel, C., Lenay, C., Marque, C., "Dynamics of Shape Recognition Through a Minimal Visuo-Tactile Sensory Substitution Interface." Third Int. Conf. On Cognitive and Neural Systems, Boston, pp.26-29, Mai 1999.

[6] Hardy, B., Ramanantsoa, M., Hanneton, S., Lenay, C., Gapenne, O., Marque, C., "Cognitive processes involved in the utilisation of a simple visuo-tactile sensory prosthesis" Sixth International Conference on Tactile Aids, Hearing Aids and Cochlear Implants (ISAC'00), Exeter, Angleterre, pp. 52-55, 2000.

[7] Hurley, S., Consciousness in Action, Cambridge (MA)/London, Harvard UP, 1998.

[8] Hutchins, E., Cognition in the wild. Cambridge, MA.: MIT Press, 1995.

[9] Lenay, C., Canu, S., Villon, P. "Technology and Perception: the Contribution of Sensory Substitution Systems", Second International Conference on Cognitive Technology, Aizu, Japan : IEEE, pp. 44-53, 1997.

[10] Lenay, C., Gapenne, O., Hanneton S., Marque C. and Genouel C., "Sensory Substitution: Limits and Perspectives", in Y. Hatwell, A. Streri & E. Gentaz (eds.), Touching for Knowing. Amsterdam/Philadelphia, John Benjamins, 275-292, 2003.

[11] Millikan, R. G. Content and vehicle. In Spatial Representation. Oxford: Blackwell, pp. 256-68, 1993.

[12] Noë, A., Action in Perception, Cambridge (MA), MIT Press, 2004.

[13] O'Regan, J.K. et Noë, A. "A sensorimotor Account of Vision and Visual Consciousness", Behavioral and Brain Sciences, 24, pp. 939-1031, 2001.

[14] Putnam, H. The meaning of 'meaning'. In Philosophical Papers, Vol. 2: Mind, Language and Reality. Cambridge University Press, 1975.

[15] Sribunruangrit, N., Marque, C., Lenay, C. Gapenne, O. and Vanhoutte, C., "Speed-accuracy tradeoff during performance of a tracking task without visual feedback" IEEE transactions on Neural Systems and Rehabilitation Engineering, 12, pp. 131-139, 2004.

[16] Thouvenin, I., Lenay, C., Cottereau, C., Aubert, D. (2003). "Design of a tactile prosthesis for blind users with a virtual reality system". In Virtual Concept, Biarritz – France, pp.199-204, 2003.

[17] Varela, F., Thompson, E. and Rosch, E., The Embodied Mind: Cognitive Science and Human Experience, MIT Press, 1991.

Walking Through a Virtual Aperture as an Assessment of Presence

Jean-Claude Lepecq* Jean-Marie Pergandi* Thelma Coyle*
Lionel Bringoux* Daniel Mestre*

(*)UMR Mouvement et Perception, Faculté des Sciences du Sport, CNRS-Université de la Méditerranée, 163 avenue de Luminy, CP 910, 13288 Marseille, France

E-mail:jean-claude.lepecq@univmed.fr, jean-marie.pergandi@ univmed.fr, thelma.coyle@ univmed.fr, lionel.bringoux@univmed.fr, daniel.mestre@univmed.fr

Abstract

This study aims at assessing behaviorally the subjective sensation of presence inside a virtual environment. The subject's task was to walk through a virtual aperture of variable width. In the case of presence, the subject's body orientation, while walking, was expected to be adapted to the width of the aperture. Most, but not all, subjects adapted their behavior to respect the virtual constraints. This behavioral adjustment is assumed to be an objective indication of presence.

1. Introduction

The concept of presence refers to the subject's belief that sensorimotor couplings occur naturally within the real world while, in fact, these couplings happen inside a Virtual Environnement (VE) [1, 3]. In order to assess presence in VE, we investigated spatio-temporal aspects of adaptive behavior, which are governed by volition and selection. In everyday life, there are well known robust adaptive behaviors, which imply both an intention to act and a selection among variations of the same act. This selection may be constrained by the relation between the body architecture and the environment. As such, evaluating presence in a virtual environment may be approached using the concept of affordances [2]. Gibson assumed that we perceive in order to operate on the environment [2]. Perception is thus designed for action. These perceivable possibilities for action were named affordances. The main hypothesis of our study is that the degree of presence in a virtual environment can be evaluated by its actual affordances for action, which can be experimentally tested.

For example, a subject may have to lengthen the step while stepping over a street gutter or to rotate the body while walking through a narrow aperture [4]. These adaptive behaviors pertain to body-scaled motor adjustments. For a street gutter of constant width indeed, the tendency to lengthen the step is more pronounced if the legs are short. Similarly, for an aperture of constant width, the tendency to rotate the body is more marked for larger shoulder widths. These body-scaled behaviors present a twofold interest. First, they are potentially elicitable within an "adequate" VE. In addition, they are objectively measurable. As such, they can provide a behavioral quantification of presence. We thus designed an experiment in which subjects had to walk through a virtual aperture whose width was variable.

2. Method

2.1. Subjects

Ten male subjects participated in the experiment, ranging in age from 21 to 30 (mean = 23.6; sd = 2.9). They had normal or corrected to normal vision. They were free from any known locomotor disorder. Their standing height ranged from 159 to 192 cm (mean = 176.2; sd = 10.5). Their shoulder width, the widest frontal body dimension, ranged from 40 to 55 cm. (mean = 45.7; sd = 3.9). Their inter-ocular distance was measured with a corneal reflection pupillometer. It ranged from 60 to 69 mm. This was taken into account in order to generate body-scaled stereoscopic images and hence individually optimize binocular vision. The stereoscopic acuity was measured using the Graded Circle Test from the RANDOT ® stereotest. It ranged from 20 to 140 seconds of arc.

The subject were equipped (Figure 1) with INFITEC ® stereo glasses (which provide stereo separation by splitting the color spectrum in two) and with reflective markers on the glasses and on both shoulders. This allowed 3D tracking of the subject's cyclopean point of gaze (for real-time updating of the visual scene) and of shoulders' positions (for offline analysis of the subject's posture) by the ART ® system. The shoulders markers were placed over the clavicule and not over the acromion to avoid subjective widening of the shoulder width. The subjects were naïve as to the purpose of the experiment.

Figure 1. 3D representation of the subject’s equipment, with stereo glasses and markers and a set of markers on each shoulder.

2.2. The virtual environment (VE)

The VE was displayed inside a particular type of CAVE ® system (Figure 2), developed at the Mediterranean Center for Virtual Reality (IFR Marey). This system consisted of 4 projection surfaces. The front, left and right walls (3 m wide * 4 m high) were back-projected acrylic screens. The rigid floor screen (3 m * 3 m) was projected from above. The top and the back faces of the cube were not projection surfaces.

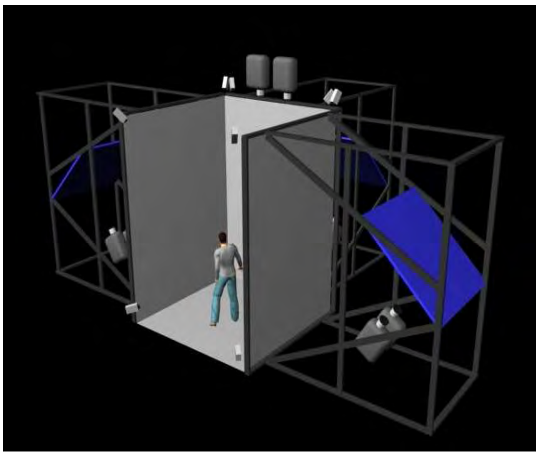


Figure 2. The Movelite system, developed in collaboration with BARCO ®. Notice the height of the screens (4 meters), providing optimal stimulation of the vertical visual field.

The VE was composed of two adjoining rooms connected via a sliding door (Figure 3). The first room was empty and was marked with a starting point (green disk displayed on the floor). The second room was furnished (in order to provide static and dynamic depth cues) and was marked with an arrival point (blue disk).

Figure 3. The virtual rooms and the sliding door

The sliding door consisted of two mobile surfaces (height = 204 cm) that could be closed or opened by lateral translation. The sliding door formed an aperture whose width was variable and ranged from 40 to 80 cm, by 5 cm steps. The subject was instructed to walk forward from the starting point straight to the arrival point, and hence had to walk through the virtual aperture (Figure 4). The locomotor trajectory was 200 cm long (110 cm before the door and 90 cm after)

Figure 4. Walking through the virtual aperture

2.3. Course of one trial

In order to optimize immersion into the virtual environment, the subjects were conducted (by the experimenter) into the VE with their eyes closed and required to open their eyes only when facing the front wall from the starting point, while the VE was displayed. In this way, they could only see the virtual environment throughout the experimental session.

The initial scene showed the sliding doors wide open. Then the doors were closed, leaving an aperture whose width was one of nine predetermined values. This closing was accompanied by a spatialized rattling sound located in front of the subject. The subject was required to walk straight from the starting point to the arrival point and to stop at this point (Figure 4). This neutral directive aimed to avoid any behavior induction by instructional semantic effects. The walking speed should be normal and comfortable. Once at the arrival point, the subject was informed that the sliding door behind him would open wide. This opening was accompanied by a spatialized rattling sound located behind the subject. When the sliding door was opened, the subject walked backwards from the arrival point to the starting point. The experimenter held the subject by the shoulders in order to guide him during this backward walk. Once at the starting point, the subject was required to precisely face the front wall. A new trial could then begin.

2.4. Experimental design

The aperture could be one of 9 widths : 40, 45, 50, 55, 60, 65, 70, 75, 80 cm. A block of trials involved 9 trials (one trial per width). Each subject performed 3 blocks (27 trials). For each block, the order of presentation of the 9 widths was randomized.

For each trial, the maximal absolute shoulder rotation was calculated from the recorded successive positions of the shoulder markers, while the subjects walked through the virtual aperture.

3. Results

The mean absolute maximum angle of shoulder rotation is plotted as a function of aperture width (Figures 5).

From the data, the population was divided into two subsets. On the one hand, two subjects did not rotate the shoulders at all while walking through the virtual aperture. At each trial, they systematically exhibited frontal walking whatever the aperture width (Figure 5 B). On the other hand, 8 subjects adapted their body orientation to the aperture width (Figure 5 A). The following analyses concern these 8 subjects only.

An ANOVA (Blocks x Apertures) was conducted on these 8 subjects. This ANOVA revealed a main effect of aperture width ($F(8, 56)=43.55$, $p < 0.001$). At first, the magnitude of body rotation follows a decreasing slope as the aperture width increases. Then, this magnitude reaches an asymptote at some baseline level at the widest apertures. This reflects the transition from body rotation to frontal walking as a function of aperture width.

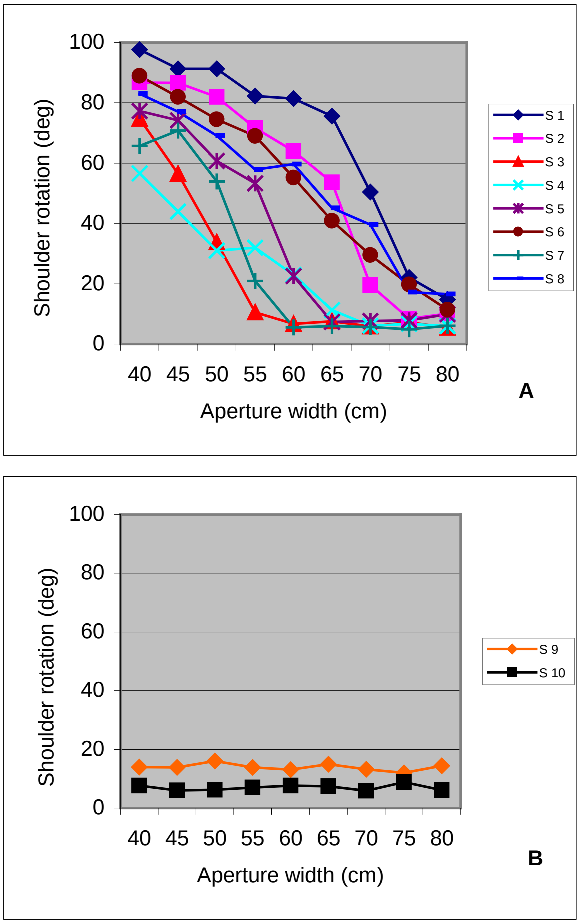


Figure 5. Individual mean absolute max angle of shoulder rotation as a function of aperture width. Eight subjects adapted their body orientation to the aperture width (5A). Two subjects exhibited frontal walking whatever the aperture width (5B).

The population of 8 subjects was then divided into two groups (small and large) based on their shoulder width. The shoulder width ranged from 40 to 45.5 cm for the small group, and from 46 to 55 cm for the large group. As expected, the large subjects tended to have greater angles of shoulder rotation than the small subjects (Figure 6 A). However, the differences between groups tended to diminish when the same shoulder rotation data are replotted against the “Aperture width / Shoulder width” (body-scaled) ratio (Figure 6 B). Thus, rescaling of the virtual aperture as a function of a relevant body characteristic eliminates group differences, suggesting that small and large subjects behave similarly relative to their own body size.

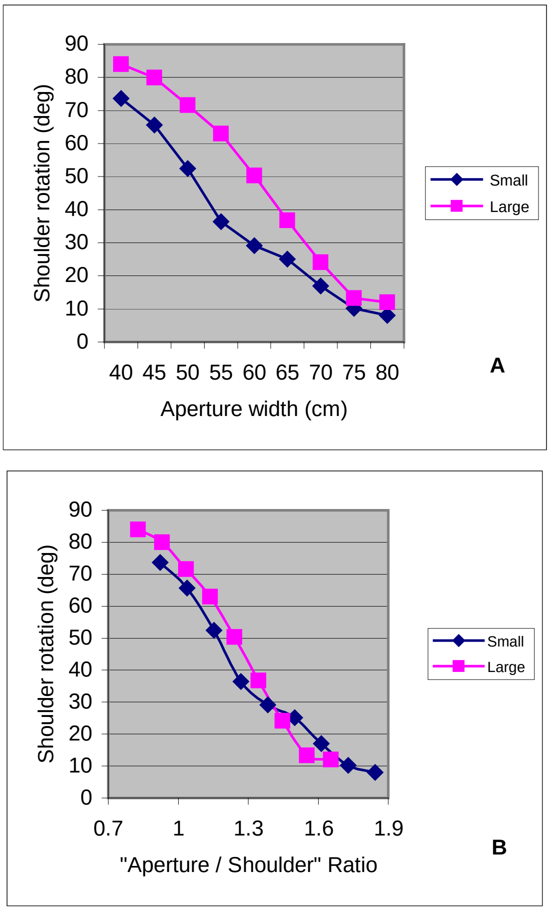


Figure 6. Average max angle of shoulder rotation, for large and small subjects, as a function of aperture width (6A) and as a function of the body-scaled ratio of aperture width divided by shoulder width (6B).

4. Discussion

The results of this study indicate that the locomotor postural patterns of subjects having to walk through a virtual aperture strongly resemble those of subjects who have to walk through a real aperture [4]. Indeed for most subjects, a behavioral transition from frontal walking to body rotation was shown as a function of the virtual aperture width. Additionnally, subjects with wider shoulder were oberved to rotate their body more than subjects with small shoulder widths. Finally, the differences between "small" and "large" subjects disappeared if the body rotation was considered with respect to a body-scaled dimensionless ratio. These two latter conclusions need further investigation with additional small and large subjects, in order to be strongly supported.

In the present study, presence is assessed by a particular motor adjustment which relates the size of a body feature (shoulder width) to the size of some characteristics in the environment (width of the door). This kind of adjustment pertains to body-scaled motor adjustment. In other words, these motor adjustments constitute some "realized affordances". According to Gibson, an affordance is an action possibility which is provided to an organism depending on boh the organism properties and the environment properties [2]. In short, the present study suggests that eliciting "acted affordances" in virtual reality research could contribute to behaviourally assess presence in virtual environments. Since any "acted affordance" implies measurable variations (e.g. magnitude of body rotation) of a given action (e.g. walking through an aperture) and that these variations depend on both some body characteristics (e.g. the shoulder width) and some virtual environment feature (e.g. the width of the aperture), it follows that any "acted affordance" can provide a sensorimotor evaluation of presence.

References

[1] Casati, R., and Pasquinelli, E. "Is the subjective feel of ''presence'' an uninteresting goal?", Journal of Visual Languages and Computing, 16, pp. 428–441, 2005

[2] Gibson, J.J. "The Ecological Approach to Visual Perception", Houghton-Mifflin, Boston, 1979.

[3] Sanchez-Vives, M.V., and Slater, M. "From presence to consciousness through virtual reality", Nature Reviews Neuroscience, 6(4), pp. 332-339, 2005.

[4] Warren, W.H., and Whang, S. "Visual Guidance of Walking Through Apertures: Body-Scaled Information for Affordances", Journal of Experimental Psychology: Human Perception and Performance, 13 (3), pp. 371-383, 1987.

Perception of Virtual Multisensory Mobile Objects Wandering around the Enactive Assumption

Annie Luciani (1), Sile O'Mondrain (2), Charlotte Magnusson (3), Jean-Loup Florens (1), Damien Couroussé (1)
(1) ICA laboratory and ACROE, INP- Grenoble, France
(2) Queen University of Belfast, Ireland
(3) CERTEC, University of LUND, Sweden

Abstract

We explore here, through three complementary experiments on virtual objects, how intimate active relations with multisensory audio-visual and haptics perceptions allow to the cognitive creation of new believable and plausible objects than can be different of the virtual ones objectively implemented. The three experiments are based on "Pebble boxes" and consist in the exploration and the manipulation of multiple moving multisensory objects (the Pebbles). They show how an inferred scene is constructed from experience, as assumed in the cognitive Enactive concept, by means of three complementary strategies: "the Emergent Exploratory Procedures (EEP), the "Dynamic Manipulation Adaptivity" (DMA), the "Adaptive Experimental Learning" (AEL). It shows also the complementarity between the ergotic and the semiotic situation on the strategies to infer a believable and plausible scene.

1. Introduction

The work presented here is at the crossing point of (1) the perception of numerous mobile objects and (2) the perception of virtual objects or real objects altered by digital processing.

On the field of perception of numerosities, a wide quantity of works exist in the numbering of perceptual stimuli: numerosity judgments of visual stimuli, counting of acoustical stimuli and more recently, of unimodal tactile and bimodal tactile/visual stimuli [1] [2]. On the field of virtual reality, most of works aim at (performing comparative experiments in order to implement realistic virtual reality platforms or (2) perceptual multisensory cues in an easier way than using real physical materials. It is a very active field new results, mainly in hapto-visual coherencies.

However, new questions are risen by the existence of Virtual objects and their multisensory manipulation. Virtual objects – i.e. things for which the behaviors is computed by digital machines, sensorially transduced by digital/analog machines, are nevertheless "real objects" as they are really presented to our perceptions and acted by our actions, and are more and more used to perform tasks in the mechanico-optical world.

The question of the perception –and more generally the identification and the appraisal – of such "strange" real objects is then a full question, rarely asked directly. Within this general question, a less addressed one is that of perception of multiple moving objects from their multisensory behaviors, including haptic, visual and auditory feedbacks to actions. The work presented here concerns such issue of perception of virtual multiple moving multisensory objects. Three aspects have been previously explored in [3]:

(1) Estimation of the number of objects: subjects are asked to assess the number of objects inside a box

(2) Sensorial preferences: the subjects were presented with different feedback sounds and asked about believability and likeability

(3) sense of control: the subjects were asked to assess how much in control of the sound he/she felt.

We aim at exploring two questions directly tied to the concept of enaction and experimented in the context of virtual or computer transformed objects:

Type 1: What is the scene (or objects) inferred from the sensory experience, i.e. from the performed action and sensory feedbacks (auditory, visual, haptic)

Type 2: Are all the sensory modalities participating to the creation of the evoked scene, and, if not, what they are and why they are not?

The paper presents three successive experiments on three different pebble boxes, with the main observed results. It concludes by a comparison between each platform under the scopes of the enactive paradigm and the ergotic / non-ergotic properties of the situation [4].

The experiments on the three platforms are performed by 10 subjects: 25 to 55 years old, 4 females and 6 males, 6 non professional musicians and 4 musicians. One of them is a native blind people.

2. Experiments on "Pebble Box 1"

2.1. Description of the platform and of the experiments

In this platform designed by [5], a real pebble box is used as a tangible interface to control synthetic sounds. The sounds produced by the real pebble collisions are picked up by a microphone and analyzed by a specific software to extract "sounds grains" events. Sound grains are used to control and trigger recorded synthesized sounds as: Bird songs, Water sounds , Crunching apple sounds, Sandpaper, …

Subjects are invited to manipulate the pebbles, freely. Three situations have been proposed with three different sounds : Birds songs, Water sounds, Crunching apple sounds. The questions were:

Q1.1. What are the scenes suggested to you?

Q1.2. Does the association of the sound and the manipulation believable, (categorized in 3 levels)?

2.2. Results

Objectively, the situations are paradoxical in the sense the palming pebbles produce sounds that are not pebble sounds.

(1) The inferred scene

In the three cases, with all the persons, an inferred - constructed - scene, that can be different from the objective one, is constructed from experience. Here are some examples extracted from subjects comments:

Birds songs

People imagine "walking on a gravel path or throwing a stone, triggering panic on birds in bushes.

Water sounds

People imagine "handling stone(s) in Water or disturbing animals (fishes) which escape".

Crunching apple sounds

People imagine "an animal within the box" and become anxious.

(2) Emergent Exploration Procedures:

The inference of a possible scene is a dynamic evolving process, in which people alternates scene assumption and exploration of ways of manipulation in order to converge to a believable inferred scene, through what we call "Emergent Exploratory procedures (EEP). For example:

When Birds songs

Gesture to walk with hands

When Water sounds

Gesture to throw pebbles in water,

Gesture of « swimming with hands »

Crunching apple sounds

Gesture to crunch something with hands,

Gesture to scratch a box in which is an animal

Prudent gestures when supposed animals or unknown living organisms nested somewhere.

(3) Believability:

First, as observed in [3], the situation when subjects hear "sound water" is the most believable. But the main non expected surprising observation, from the observers and from the subjects, done in such type of experiment, when the exploration is free , and when analyzing spontaneous and free comments, is that there are no inferred scenes that could be totally un - believable. Even in the cases looking very far than a real situation (such as the bird or the apple crunching sounds), subjects are likely surprised, but « nicely surprised » (See photograph in Figure 2. The situation did not seem to them completely impossible. They modify their manipulation and their interpretation to make as best as possible the situation believable.

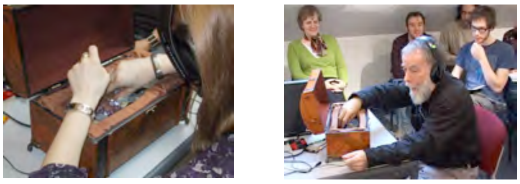


Figure 1. Left : "Pebble Box 1"
Right : An experiment: "surprising but believable";

3. Experiments on "Pebble Box 2"

3.1. Description of the platform and of the experiments

The platform is the same that used in [6]. People handle 10 physically-based 3D cubes or spheres, simulated by the Open Dynamics Engine software in squared box (Figure 3). They stir up by means of a Phantom Omni device. Sounds of collisions are triggered by collision detection algorithm. There are different possibilities for the visualization:

- The ten objects are visually represented or not
- The manipulator is visible or not
- Pebbles could be cubes or spheres.
- Visual size can be different to haptic size.

Figure 2. Pebble Box 2: VR Haptic manipulation & "Changing the visualization"

3.2. Results

Globally, and similarly than in Pebble Box 1, people try to infer a believable scene, if possible from all the multisensory feedbacks. But differently than in Pebble Box 1, and if they cannot, they are led to elude some modalities, the eluded sensory feedback being not always the same for all the experiments.

(1) When no visualization, people trend to infer a phenomenon rather than « clearly cut objects », (for example "pebbles"). For example

- In addition to clearly cut objects feeling, people talk about « force field », « magnetic field », paste, medium resistance, grain in paste, etc…

- When visual objects are smaller than their physical radius, the physical inferred objects are supposed surrounded by a transparent shell or extended by a force field.

(2) Subjects change the way they manipulate when they have visual perception of themselves. We observe here the processes of Emergent Exploratory procedures (EEP). For example:

- *They explore the whole space more when they have visual feedback of themselves than without*
- *They attempt to create the conditions allowing them to explore the shape of the supposed objects.*

4. Experiments on "Pebble Box 3"

4.1. Description of the platform and of the experiments

The Pebble box 3 is a 2D Virtual Pebble Box composed of a circular box containing 8 mobile masses more or less rigid, in interaction of collisions more or less visco-elastic and with 1 more or less rigid mass (the manipulator) controlled by the ERGOS haptic high quality device.

Figure 3. Pebble Box 2: "Changing the matter & the visualization" & VR Haptic manipulation

By changing the physical parameters of the interactions between each pairs of pebble-masses and also between each pebble-mass and the haptic stick-mass, we modeled the matter changing. Following a first serie of experiences done [7], four well-categorized cases have been chosen:

	1	2	3	4
Pebbles Rigidity	high	low	high	high
Stick-pebble rigidity	high	low	high	high
Viscosity	medium	medium	Low	Low
Pebble size	Big	medium	medium	small

Two visualizations were used: (1) a ball like visualization; (2) a blurred medium-like visualization

The sounds are the sounds produced by the simulation of the pebbles at acoustical frame rate (44KHz). When colliding, the two pebbles are vibrating producing one sound each, as in the real mechanical world. The sound signals depend on the physics of the collision (matter of the colliding objects, strength of the hit, velocities, etc.). This simulation is objectively a very realistic physical simulation of identical pebbles pushed with another object.

4.2. Results

Similarly than in Pebble Box 2, people try to infer a believable scene, if possible from all the multisensory feedbacks, and if not, they are led to elude some sensory feedbacks. As previously, the inferred scene is not necessarily coherent with the objective simulated one, confirming here too a creation of a plausible scene from the sensori-motor experience.

(1) The value of physical material interactions leads to infer two types of different categories of scene not necessarily similar to the objective one. For example:

- *People feel a kind of « medium », « paste, « force field », « cotton », etc. when grains are in soft colliding interactions*
- *People feel clearly-cut objects but not necessary all of them or of the same size.*
- *When the sound or the vision are not consistent, they are preferably eluded.*

(2) The dynamic of the coupling of the manipulation, mainly the intensity of grasping done by the performer depends of the implemented scene. People adapt his own dynamic and the dynamic of coupling to the physical constitution of the manipulate object. We are confronted here to a kind of Dynamic Manipulation Adaptivity (DMA). For example:

- *When objects are in a strong rigid interaction, people grasp strongly the device and act (presses, moves) with high energy.*
- *When objects are in soft elastic interaction, or when they are small, people manipulate delicately for example by grasping the stick with fingers.*

(3) The refinement of the exploration increases along the experience and the scene inferred change progressively. We are confronted here to a a kind of Adaptive Experimental Learning (AEL). For example:

- *When the simulated matter is very soft, people start with feeling nothing and progressively tend to feel a type of « resistant or viscous » field, or field+lumps*
- *When the simulated matter is Very rigid and the objects very big, without visualization, People start with feeling « one big object », explore its shape and progressively discover eventually the others, that are imagined smaller.*

5. Comparison

5.1. Enactive experience

In all the experiments, with all the subjects, an inferred scene is constructed from the sensori-motor experience. This inferred (constructed) scene can be different from the objective scene. We can talk about "a created scene". There is more a cognitive creation of objects than an identification of objects.

Using Virtual objects to explore how objects as cognitive categories emerge, we can say that these observations, in a free experimental context, fall on the scope of the enactive paradigm: *"...cognition is not the representation of a pregiven world by a pregiven mind but is rather the enactment of a world and a mind on the basis of a history of the variety of actions that a being in the world performs."* [8].

During the process of inference of a plausible scene, subjects developed three complementary strategies:

• Emergent Exploratory Procedures (EEP) in which subjects are seeking for on-line exploratory procedures in concordance with assumption on the possible scene suggested by actions and sensory feedbacks.

• Dynamic Manipulation Adaptivity (DMA) in which the subjects adapt very quickly and "on the fly" his manual dynamic performance to the sensory feedbacks and the supposed felt objects.

• Adaptive Experimental Learning (AEL): We observe that subjects learn very quickly, dynamically and "on the fly", what are the best manipulations. This is a derivative property of intimacy and embodiment: subjects are "with" the object.

5.2. Ergotic / non-ergotic interactions

One of the reason for which we performed the experiments on the three "Pebble Boxes" is their complementarity regarding the physical consistency between gestures and sensory feedbacks. An efficient typology for that is the typology introduced by Cadoz [4], who distinguishes ergotic situations, and non – ergotic ones, i.e. purely epistemic and/or semiotic. In ergotic situations, there is a physical energetic coherence from action to the sound and the image that are so produced. This property is a fundamental property to support intimacy and embodiment [4]. In non ergotic interactive situations, such as the mouse – visual one, sign language, musical conductor control, etc., the information exchanged are purely semiotic and epistemic. We noticed that:

• In pebble box 1, the relation between the gestures and the auditory sensory feedback is purely semiotic / epistemic.

• In pebble box 2, the relation between action and visual feedback is totally managed by physical modeling and so it is an ergotic relation. Conversely the sounds are triggered from a signal event and consequently, it is a non-ergotic situation.

• In pebble box 3, the relation between action and all the visual feedbacks (visual and auditory) is totally managed by physical modeling and so it is a multisensory ergotic relation.

Two remarkable observations can be extracted from the performances:

1. More the situation is ergotic, more are the effects of dynamic adaptation and dynamic on-line learning. Less are believability of the whole situation, in the sense that some modalities are sometimes eluded to conclude to a plausible effect. This is very noticeable in pebble box 3.

2. Less the situation is ergotic, more are the importance of the emergent exploratory procedures and the metaphoric prospect. In addition, most of the scenes are believable in accordance of the whole sensory returns.

6. Conclusions

We conducted here observations in order to go initiate new research on what could be the perception of virtual objects, assuming that virtual objects have also to be considered as real objects, while they can be explored sensorially in a stable context. Such researches will be complementary to others aiming at whether using virtual objects as versatile experimental settings to go forward in the understanding of the human perception or using virtual reality platforms for training.

Differently than in these others experiments, usually based on very precise questions (subjective or quantitative), the exploration we propose to the subjects are totally free, as if they were in front a really new object he / she is discovering. We can see that Virtual Reality could be "a laboratory to capture how human construct what an object is". Thus, the method is to analyze "natural human behaviors and spontaneous and free comments", when they are discovering through their action and perception, without any other preliminary comments form the experimenter, such "strange new things".

References

[1] Gallace A, Tan H Z, Spence C, 2006, "Numerosity judgments for tactile stimuli distributed over the body surface" Perception 35(2) .

[2] Gallace A, Tan H Z, Spence C,, "Multisensory numerosity judgments for visual and tactile stimuli," to appear in Perception & Psychophysics, Vol. 69, No. 4, 2007.

[3] G. Essl, C. Magnusson, J. Eriksson, Sile O'Modhrain. Towards evaluation of performance, control of preference in physical and virtual sensorimotor integration. Enactive 05. 2nd Conference on Enactive Interfaces. Genoa. November 2005.

[4] Cadoz C., Wanderley M. (2000). "Gesture-Music". In "Trends in Gestural Control of Music", M. Wanderley and M. Battier, eds, -2000

[5] O'Modhrain M. S. , Essl G. "PebbleBox and CrumbleBag: Tactile Interfaces for Granular Synthesis". NIME 2004: 74-79.

[6] Magnusson C., Luciani A. Couroussé D., Davies R. Florens J.L.. « Preliminary test in a complex virtual dynamic haptic audio environment ». 2nd Enactive Workshop May 25-27 2006 - McGill University, Canada

[7] Maxime Houot, Annie Luciani, Jean-Loup Florens. "Perception of multiple moving objects through multisensory-haptic interaction: Is haptic so evident for physical object perception?". Proc. of Enactive06 Conference. Montpellier. Nov. 2006

[8] Varela, Thompson, Rosch. The Embodied Mind . 1991.

Acknowledgements

This research work has been performed with the support of the FP6 NoE Enactives Interfaces and with the French Ministry of Culture.

Drawing and guiding gestures in a mathematical task using the AHEAD application

Charlotte Magnusson[*] Kirsten Rassmus-Gröhn[*] Håkan Eftring[*]
(*)Department of Design Sciences, Lund University, Sweden
E-mail: charlotte.magnusson@certec.lth.se, kirre@certec.lth.se, hakan.eftring@certec.lth.se

Abstract

This paper presents partial results from an evaluation done in practical school work of the audio haptic drawing editor and explorer (AHEAD). In this paper we focus on a mathematical design and reproduce group work task performed by two pupils (one sighted and one visually impaired pupil). Apart from the fact that the pupils were able to do the task, tentative results show some differences in the usage of the program functionality while no significant differences can be seen in the times spent drawing. We also present some discussion of the guiding gesture used by the mouse user to help the PHANToM user understand the drawn image.

1. Introduction

Getting access to 2D graphics is still a large problem for users that are severely visually impaired. Using a haptic display in combination with audio feedback is one way to enable access. General guidelines to create and develop haptic applications and models are collected in [1]. Applications making practical use of non-spoken audio and force-feedback haptics for visually impaired people are e.g. applications supporting mathematical display [2], [3] & [4], games [5-7] and audio-haptic maps [5;8]. As described in [9] and [10], there are indeed people who are blind who have an interest in hand drawing. In [11], a CAD application is presented that enables users to create drawings with the help of audio and keyboard. In [12], a study on a haptic drawing and painting program is presented.

2. The AHEAD application

The AHEAD audio haptic drawing editor and explorer application (see figure 1) is well described in [13] and we give only a short summary here. The virtual environment consists of a virtual sheet of paper that is oriented in the vertical plane (standing up). The application can be used in two different modes: one for editing and one for exploring relief drawings. The application supports two users, one mouse user and one PHANToM user. Both users can draw, edit and explore on the virtual paper. There is also a guiding function where either user can guide the pointer of the other user. The PHANToM user can pull the mouse to the PHANToM position and the mouse user can drag the PHANToM to the mouse position (in this case the PHANToM user experiences a constant force which drags the stylus towards the mouse position).

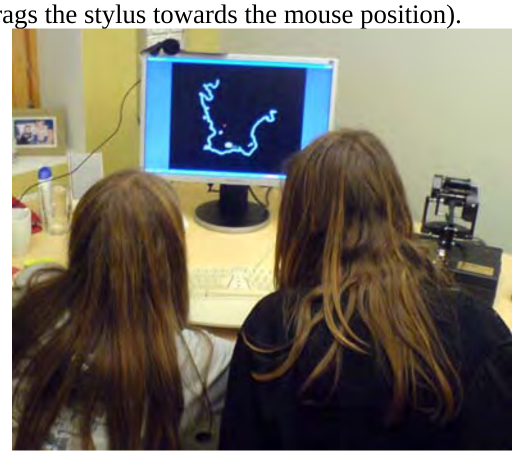


Figure 1. Two pupils using the AHEAD application.

3. The mathematical design and reproduce group work task

The test was conducted in a ninth grade class in the subject mathematics. The learning task in the particular lesson was to practice using geometrical mathematical language, i.e. words like “rectangle”, “sphere”, “angle” and “diagonal” to describe a composite geometry figure to a fellow pupil. The sighted pupils were instructed to use paper, pencil and a ruler, and the visually impaired pupil and a fellow pupil were instructed to use the AHEAD application. The test setup consisted of a laptop running the program and acting as screen and keyboard for the sighted pupil who also was using a mouse for input. The visually impaired pupil had a separate keyboard attached to the same computer, a screen, headphones and the PHANToM OMNI. Half of each screen was blinded by a piece of cardboard to prevent the pupils from seeing the drawing the other person made.

The AHEAD application was loaded with a file with a subtle grid in positive relief, and a middle line with the spoken caption "Stop, middle line" (see figure 2). The pupils were supposed to use one part of the virtual paper/screen each to draw on and the middle line was not to be crossed until the last phase of the task. There were three parts to the task; first, one pupil would design a composite figure in the drawing application (without showing it to the other pupil); second, the same pupil would describe the figure to the other pupil who would try to make a copy based on the description; third, the pupils would together compare the copy to the original figure.

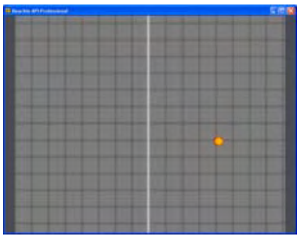


Figure 2. The empty grid for the mathematic task..

4. Results

If we start by looking at the resulting images (figure 3) we see that the pupils have succeeded with the task – i.e both have managed to understand the instructions and show this understanding by drawing a reasonable replica of the original (the drawings are not perfect, but they catch the essence of the design).

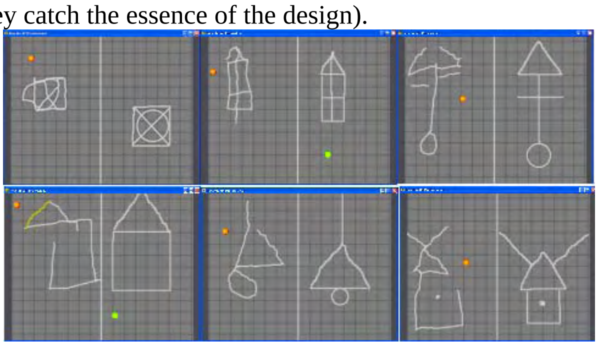


Figure 3. The pairs of drawings generated. In the top row, the mouse user (to the right) designed the original, while in the bottom row it was the PHANToM user that did this. The time order in these images are task 1,3 and 5 in the top row and 2,4 and 6 in the bottom row.

Since the AHEAD application logged both sampled PHANToM and mouse positions, together with clicks and keyboard commands it is possible to look a little more in detail at the process behind these drawings.

If we start by looking at the PHANToM movements (figure 4) we see that the PHANToM movements are mainly concentrated where they should be, indicating that the user does not have problems with disorientation, something which was supported by observations made during the test – the visually impaired PHANToM user seemed quite well aware of where things were, and did not appear to have any major problems getting lost within the workspace (despite the absence of any limiting box in the current version of the AHEAD application). One problem that did occur was that the user on one occasion accidentally erased the wrong line – something which indicates that more feedback is needed for this type of operations – but apart from this no major problems were experienced. In the top row one can see a few points in the wrong side of the workspace. Since these are not present in the lower row we interpret these to be mainly "resting points" – the user is not really using the PHANToM, but is instead waiting for the mouse user to complete the design.

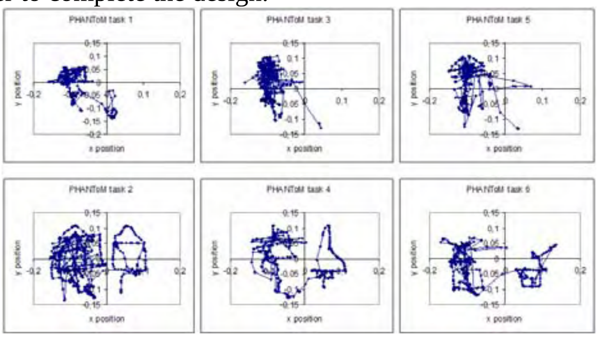


Figure 4. Sampled PHANToM x-y positions for the different tasks.

In the bottom row one can also see the guided exploration of the mouse generated drawing to the right. It is clear that this guided motion is quite different from the exploration on the left side – the motion on the left side shows more scanning behavior. The PHANToM positions can be compared to the sampled mouse movements shown in figure 5, where the mouse is used not so much for exploration, but mainly for drawing/guiding and quick pointing gestures. The fact that the faster mouse motion is more spread out over the workspace can be partly explained by the fact that the mouse also is used for interaction with the file open dialogs – and of course by the fact that it may be moved unintentionally while the mouse user is waiting for the PHANToM user. For the mouse it is quite clear that the drawing motions are slower than the other gestures (dots closer in the diagrams), but for the PHANToM this is not so obvious. There are some fast explorative moves, but also slower and more detailed ones.

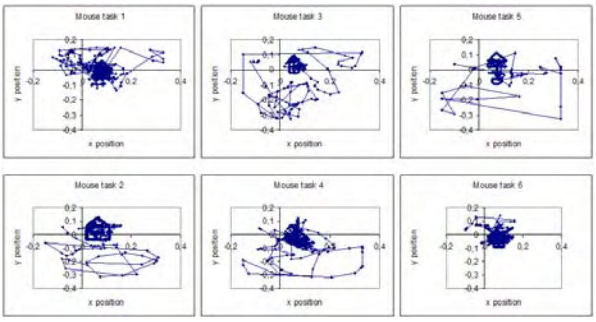


Figure 5. Sampled mouse x-y positions for the different tasks.

If we look at the amount of erasing actions performed (table 1) we can see a clear learning effect. Both users start by having to erase several times, but after a while they are more familiar with both the tool and the exercise, and no longer need to erase any lines.

Table 1. Nr of erasing actions. The user drawing the original is indicated by *.

Task	PHANToM	Mouse
1	0	6*
2	3*	1
3	3	0*
4	0*	0
5	0	0*
6	0*	0

Looking closer at the drawing gestures for task 1 we can see that the sighted user has experimented quite a lot by drawing and erasing, and this user also used the moving functions to move the lines to the right positions, as well as the shape functions to change lines into straight lines, circles or rectangles. The visually impaired user did not use any of these functions and drew everything where it should be from the beginning (and did not bother to make perfect circles, lines or rectangles).

In task 2 the visually impaired user actually succeeded quite well drawing the figure initially, but then accidentally drew a small line and when attempting to erase this unwanted line instead erased the larger drawing. Because of this the other lines were also erased and the user started from the beginning again. The second time this user had more problems positioning the lines correctly, but it is still possible to see the intended figure (see figure 3, bottom left).

In task 3 the sighted mouse user immediately gets things the way she wants them, while the visually impaired pupil gets it a bit wrong initially, and instead of using the keys for moving the lines, the lines in the erroneous position are erased and new ones are drawn.

For the tasks 4, 5 and 6 the only change made is by the sighted user who uses the straight line function once in task 4 and 5 each. Otherwise the drawings are made the way they are shown in figure 3.

The time spent drawing can be seen in table 2. If we look at the all six tasks and check for differences between the time spent drawing, or the average drawing time per line no significant differences can be found between the PHANToM and the mouse (t-test, t=0,98 and 0,95 respectively). Despite this, some more qualitative observations can still be made. If we look at the total time spent drawing we see that for task 1 and task 2 (the first time each user takes on the different roles in the task) the user designing the original spends longer time drawing than the person drawing the copy. For the following tasks the opposite is true, i.e making the copy takes longer. This should be expected since making the copy involves more talking – the user making the copy has to get the spoken instructions. That things are different the first time is also quite understandable – the pupils have to get used to both the equipment and the exercise. No such differences can be seen for the average time per line, which may indicate that the difference is not so much due to the drawing as it is to the other activities in the tasks. Disregarding the first two tasks it seems that the ratio between the total drawing time of the designer and the copier is smaller when the mouse user is the designer compared to when the PHANToM user is the designer. Since the time spent drawing different lines does not differ this should be due to some other part of the task – the spoken part of the exercise and/or the cognitive/explorative part. Closer examination of the time span between drawing the lines does not reveal significant differences between the PHANToM and the mouse (t=0.76) although one can see that the PHANToM user performs more consistently than the mouse user.

Table 2. The time in seconds spent drawing for the different tasks. For task 1,3 and 5 the mouse user is the one that designs the original, while for the tasks 2,4 and 6 it is the PHANToM user that does this.

	PHANToM		Mouse		Ratio, total
Task	total	average per line	total	average per line	Designer/ copier
1	79	7	164	4	2,1
3	141	5	47	6	0,3
5	151	4	42	7	0,3
2	194	6	43	10	4,5
4	42	5	77	6	0,5
6	86	8	134	8	0,6

Another interesting issue is how the guidance was used. When the sighted user drew a copy, this was followed by mouse guiding to present the result to the visually impaired pupil (in the opposite condition there was no guiding – the sighted user just checked the results visually). If we look at the three cases when mouse guidance was used we see that the mouse user tends to start the gesture a bit away (nearer to the position of the PHANToM user) and then guide the PHANToM towards and around the drawing. This strategy appeared to work well – particularly if we compare it to what has happened in other tests where the mouse user sometimes just put the mouse at the desired position and tried to drag the PHANToM from there (no real guiding motion).

5. Conclusion

In this test we find the AHEAD application useful and the students were able to perform the task. The editing functions (apart from erasing) such as moving and changing curves to circles, rectangles or straight lines were only used by the sighted user. For the mouse we see a difference between slower drawing gestures and quicker pointing/moving gestures. For the PHANToM this difference is not so clear – there appears to be both slow and fast exploratory movements. Surprisingly enough there is no significant differences between the time spent drawing for the PHANToM and the mouse, apart from a tendency for the ratio between the time spent drawing for designer and copier to be smaller when the mouse user is the designer (disregarding the first time each user does the designing). Due to the low number of trials and the nature of the task it is hard to say anything definite about this, but since the time spent drawing individual lines appear to be quite similar it is possible the difference is due to the verbal/cognitive part of the task. We note that our mouse user uses a guiding strategy that works quite well – the mouse user “catches” the PHANToM away from the line and guides it towards and around the drawing. Finally, we see that the function where the PHANToM user could drag the mouse to the PHANToM position was not used.

6. Acknowledgements

This work is funded by the EU/ IST eInclusion projects MICOLE (IST-2003-511592 STP) and ENACTIVE (IST-2002-002114).

The authors are grateful to the EU for the support given in carrying out these activities.

References

[1] Sjöström, Calle, "Non-Visual Haptic Interaction Design - Guidelines and Applications." Ph D Dept. for Design Sciences, Lund University, Faculty of Engineering, 2002.

[2] Yu, W. and Brewster, S. A. Comparing Two Haptic Interfaces for Multimodal Graph Rendering. 2002. Florida, USA.

[3] Yu, W., Kangas, K., and Brewster, S., "Web-based haptic applications for blind people to create virtual graphs," Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2003.HAPTICS 2003.Proceedings.11th Symposium on, pp. 318-325, 2003.

[4] Bussell, L. Touch Tiles: Elementary Geometry Software with a Haptic and Auditory Interface for Visually Impaired Children. 512-515. 2006. 2003.

[5] Iglesias, R., Casado, S., Gutierrez, T., Barbero, J. I., Avizzano, C. A., Marcheschi, S., and Bergamasco, M., "Computer graphics access for blind people through a haptic and audio virtual environment," Haptic, Audio and Visual Environments and Their Applications, 2004.HAVE 2004.Proceedings.The 3rd IEEE International Workshop on, pp. 13-18, 2004.

[6] Magnusson, C., Rassmus-Gröhn, K., Sjöström, C., and Danielsson, H. Navigation and Recognition in Complex Haptic Virtual Environments - Reports from an Extensive Study with Blind Users. Wall, S. A., Riedel, B., Crossan, A., and McGee, M. R. 2002. Edinburgh, UK.

[7] Magnusson, C. and Rassmus-Gröhn, K. Audio haptic tools for navigation in non visual environments. 2005. 11-17-0050.

[8] Magnusson, C. and Rassmus-Gröhn, K., "A Virtual Traffic Environment for People with Visual Impairments," Visual Impairment Research, vol. 7, no. 1, pp. 1-12, 2005.

[9] Kennedy, J. M., Drawing and the Blind - Pictures to Touch New Haven and London: Yale University Press, 1993.

[10] Art Beyond Sight Art Education for the Blind, Inc. and AFB Press, 2003.

[11] Kamel, Hesham M., "The Integrated Communication 2 Draw (IC2D)." Ph D Electrical Engineering and Computer Sciences Department, University of California, 2003.

[12] Hansson, Christine, "Haptic Drawing Program." Master Design Sciences, Lund University, Faculty of Engineering, 2003.

[13] Rassmus-Gröhn, K., Magnusson, C., and Eftring, H. Iterative Design of an Audio-haptic Drawing Application. 2007. CHI 2007. 4-28-0070.

Force design for memory aids in haptic environments

Charlotte Magnusson* Kirsten Rassmus-Gröhn*
(*)Department of Design Sciences, Lund University, Sweden
E-mail: charlotte.magnusson@certec.lth.se, kirre@certec.lth.se

Abstract

This paper presents a study designed to investigate the usefulness of different force designs for haptic memory aids (beacons). The results lead us to recommend constant type forces, possibly with some 1/sqrt(r) like snapping behavior in the vicinity of a beacon point to make sure you actually reach it. This snapping has to be weighed against the possible disadvantage of interfering with user exploration close to the points. We also recommend that the user should be able to adjust the strength of this type of force.

1. Introduction

With one point haptic interaction in a non-visual setting, it is easy to miss objects or get lost in haptic space [1]. Some navigational tools have been suggested, such as “magnets”, “crosses” (allowing the user to feel if he or she is aligned with an object) or a “ball” (to feel things from a distance) [2]. The attractive force in particular has been used and found to be helpful in many circumstances (e.g. [3, 4] and is included as a standard tool in the current OpenHaptics software from SensAble). For graph exploration, Roberts et al. [5] and more recently Pok-luda and Sochor [6] presented different versions of guided tours, while Wall and Brewster [7] tested the use of external memory aids, so called “beacons”, which the users could place on a surface and which then could be activated to drag the user back to this particular location. Text labels have been used extensively to help users obtain an overview of maps [8] or traffic environments, for example [9].

Other suggested ways to help the user with navigation/learning are automatic guid-ing constraints, referred to as “fixtures”, which have been used for tele-operation, shared control tasks, tracking and training, often in a medi-cal context [10], or to have the user cancel forces generated by the haptic device [11].

In our previous work on navigational tools [12],[13] we had confirmed the results obtained by Wall & Brewster [7] which was that attractive forces can be useful for helping users to locate targets when using the PHANToM. We had also noted that the type of attractive force used could influence the results and we decided to perform a small study to compare different types of attractive forces. The results of this study are presented below.

2. Implementation

The current study is motivated by the fact that to make effective attractive force beacons for a PHANToM environment, one needs to know more about how different users are able to work with different types of attractive forces. Thus we decided to test forces that increased towards the target, forces that were kept constant over distance and forces that increased as you move away from the target. Our previous results indicated that an 1/r force probably was too strong at close distances [12], [13], while results from Wall et al [7] indicated that the linear force produced too strong forces at longer distances.

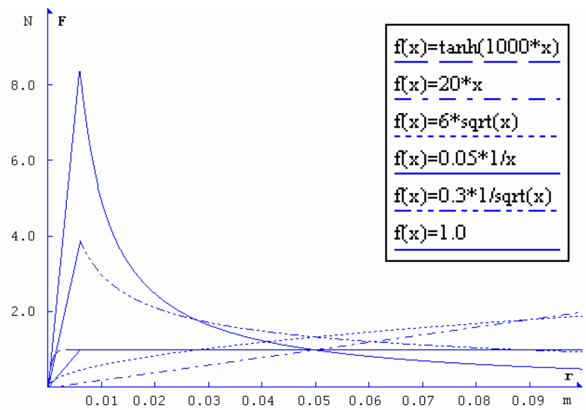


Figure 1: The radial dependence of the different forces used. The tanh and constant forces differ only at very small distances.

Because of this we decided to keep these two types as “endpoints” of our scale and add other forces in

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

between. The six different radial dependencies we used were: constant force, tanh(r), 1/r, 1/sqrt(r), sqrt(r) and r.

To avoid vibrations on the beacon position, the forces that did not tend to zero at small distances had a short linear part for very small distances (inside a radius of 0.006m). This linear part was attached so that the force function was continuous throughout the whole space (although the derivative would be discontinuous at the breakpoint). The forces were adjusted by hand to feel roughly the same at medium distance (0.05 m). The radial dependence of the forces is illustrated in figure 1. The forces were always directed along $\hat{e}_r$ towards the target.

For this test the PHANToM 1.0 premium model was used, since it has more precise force rendering. It should be noted that the strength of the forces needs to be adjusted if another PHANToM model is used. The test environment was the haptic-audio drawing program developed at Certec, Department of Design Sciences, Lund University.

3. Test design

The test consisted of 5 tasks. In task 1 the user was asked to rate how well they liked the different forces as they were being held to a point in space by the force. In task 2 the user was guided to three different points quite far apart by use of the forces and the user was asked to rate the forces for this type of task. The user initiated beacon changes himself/herself. The task 3 was the same as the task 2 apart from the fact that the test leader initiated the change of beacon. In task 4 the user was guided between two nearby points by the use of the forces and rated the forces also for this case. The user initiated a change of beacon himself/herself.

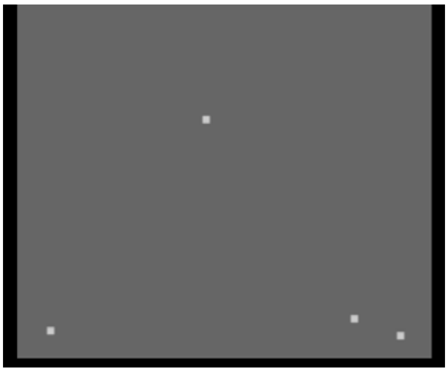


Figure 2: The points used for the test. The top point was used for tasks 1 and 5 while the three points closest to the borders were used for tasks 2 and 3. The two points in the corner was used for task 4.

In task 5 the user was asked to close his/her eyes and to draw a circle starting from the beacon point and then using the beacon force to close the circle, and then to rate the forces for this type of task.

The order the forces was presented was the same for all tasks for one user, but the order was changed between users to avoid learning effects. Figure 2 shows the points used for the five different tasks.

4. Test results

Fourteen participants between the ages of 10 and 73 did this test. Based on the assumption that this kind of basic interaction will provide reasonably similar results for blind and sighted participants, due to the limited availability of blind test persons we did this test with sighted users.

The main result from this test was the ratings of the different force designs. Qualitative observations of the interaction were also made during the tests.

The average ratings for different forces and different tasks are shown in figure 3.

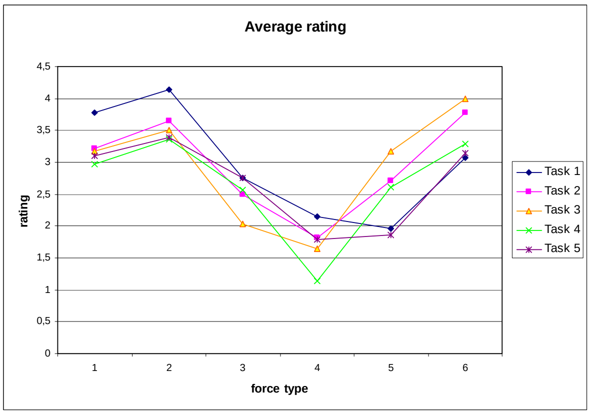


Figure 3: The average ratings for different forces and different tasks. The forces were 1. constant force, 2. tanh(r) type force, 3. sqrt(r) type force, 4. linear force (r type force) , 5. 1/r type force and 6. 1/sqrt(r) type force.

The results were analyzed using five different analyses of variance (ANOVA) with the rating as dependent variable. The independent within-group variable was the six different force designs. Post hoc tests were carried out using the Tukey test. The significance level was set to 0.05 throughout the analyses.

The ANOVA for the rating showed significant differences for all five tasks ($F(5,65)=11.9$, 14.0, 16.9, 9.2 and 7.8, $p<0.05$).

For task 1 the post hoc test showed that the tanh(r) type force was rated significantly higher than all other forces except the constant force ($Q(6,65)= 5.5$, 7.9, 8.6 and 4.2, $p<0.05$). The constant force was rated significantly higher than the linear force ($Q(6,65= 6.5$, $p<0.05$) and the 1/r type force ($Q(6,65)= 7.2$, $p<0.05$).

Finally the 1/sqrt(r) type force was rated significantly higher than the 1/r type force ($Q(6,65)= 4.4$, $p<0.05$).

For task 2 the post hoc test showed all forces except the sqrt(r) force to be rated significantly higher than the linear force ($Q(6,65)= 7.0, 9.1, 4.5, 9.9$, $p<0.05$). The tanh(r) force was also rated significantly higher than the sqrt(r) force ($Q(6,65)= 5.7$, $p<0.05$) and the 1/r force ($Q(6,65)= 4.7$, $p<0.05$). Also the 1/sqrt(r) force was rated significantly higher compared to the same set of forces ($Q(6,65)= 6.5, 5.4$, $p<0.05$).

For task 3 the constant, tanh(r), 1/r and 1/sqrt(r) forces are all rated significantly higher than the sqrt(r) and linear forces ($Q(6,65)= 5.2, 6.7, 5.2, 9.0, 7.0, 8.5, 7.0, 10.8$, $p<0.05$).

For the task 4 all forces are rated significantly higher than the linear force ($Q(6,65)= 6.8, 8.3, 5.4, 5.5, 8.0$, $p<0.05$).

Finally, for task 5 the constant force and the tanh(r) force were rated significantly higher than the linear force and the 1/r force ($Q(6,65)= 5.3, 5.1, 6.5, 6.2$, $p<0.05$). Also the 1/sqrt(r) type force was rated significantly higher than these two forces ($Q(6,65)= 5.5, 5.2$, $p<0.05$).

These results are confirmed by the qualitative observations for the different forces showed particular problems with both the linear and the sqrt(r) type force since both these forces tended to be too strong at long distances while at the same time not being able to guide the user all the way to the target, since they became too weak at short distances. Figure 4 shows a drawing which illustrates this point.

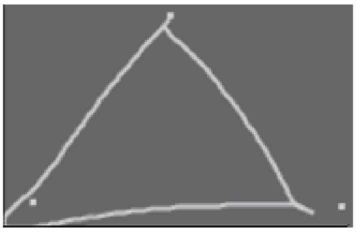


Figure 4. A drawing showing a result from task 2 with the linear force. One observes that this force was not able to guide the user all the way to the different points (the white dots in the picture).

It was also observed that the 1/r type force tended to be too strong at short distances, and the users found it much easier to interact with the 1/sqrt(r) force behavior, particularly for the task where they were supposed to draw a circle – it took quite a lot of force to “break free” from the beacon. Despite this they were still able to draw quite nice (and closed) figures as can be seen to the left in figure 5. To the right is a figure showing the circle resulting from the linear force. Also for this type of task the fact that the force gets too weak at short distances generates problems.

Figure 5. To the left a circle drawn from a beacon using the 1/r type force. To the right a circle drawn from a beacon using a linear force.

During the test we also noted that users needed different strengths of the forces, which indicate that it should be possible for the user to adjust the strength of the force.

Finally we did not note any real difference between the smooth tanh(r) force and the non-smooth constant force, although some users rated them slightly differently.

5. Discussion and conclusion

The results in the previous section show that the constant type forces (the constant force and the tanh(r) force) generally do well. Also the 1/sqrt(r) type force gets good marks by the users. The ratings are supported by the qualitative observations made during the test. It is interesting to note that this shows that users are able to deal with very much larger forces towards the end of a movement, compared to what can works at the start. Another way to say the same thing is that when forces increase continuously during the motion they are much easier to deal with compared to when sudden changes in the forces experienced.

To conclude, the advantages and disadvantages for the investigated forces can be summarized as follows:

Constant force and tanh(r) type force: Easy to predict. Takes the user to the targets while not interfering too much with user exploration. Does not snap to the targets which can cause overshooting before the user actually gets to the target.

Sqrt(r) type force: Does not interfere so much with user exploration, but does not reliably take the user to the targets. Increasing the strength to improve this makes the force too strong at large distances.

Linear force: Similar to the sqrt(r) force only more pronounced. Does not work at all at short distances while being very strong at large distances.

1/r type force: Gets the user reliably to the points. Outside the vicinity of the point it does not interfere much with user exploration. Too weak at large distances and very hard to pull free from a point.

1/sqrt(r) type force: Gets the user reliably to the points. Outside the vicinity of the point it does not interfere much with user exploration. Less hard to pull free from a point, although some users thought it was still a bit difficult to pull free from a point.

To summarize the above, we can say that for the type of tasks studied, users like forces which do not interfer too much with exploration and that, depending on the task, some short distance snap-to-point behaviour is useful. In later applications we have tested these types of forces also for mobile beacons (such as when the PHANToM is dragged by the mouse in the haptic-audio drawing program) and they have been shown to work well. In the current implementation of our program we use the 1/sqrt(r) type force both for stationary and mobile beacons.

Thus we recommend constant type forces, possibly with some 1/sqrt(r) like snapping behavior in the vicinity of a point to make sure you actually reach it. This snapping has to be weighed against the possible disadvantage of interfering with user exploration close to the points. We also recommend that the user should be able to adjust the strength of this type of force.

6. Acknowledgements

During this work we have discussed with Steve Wall & Johan Kildal at UGLAS, and we are grateful for their comments & suggestions.

This work is funded by the EU/ IST eInclusion projects MICOLE (IST-2003-511592 STP) and ENACTIVE (IST-2002-002114).

The authors are grateful to the EU for the support given in carrying out these activities.

References

[1] Colwell, C., Petrie, H., Hardwick A. and Furner, S. (1998) Haptic Virtual Reality for Blind Computer Users, Proceedings of the ACM Conference on Assistive Technology - ASSETS 98, pp 92-99

[2] Sjöström, C. (1999) “The IT Potential of Haptics – touch access for people with disabilities”, Licentiate thesis, http://www.certec.lth.se/doc/touchaccess/

[3] Wall, S.A., Paynter, K., Shillito, A.M, Wright, M., Scali, S. (2002) The Effect of Haptic Feedback and Stereo Graphics in a 3D Target Acquision Task, In Proc. Eurohaptics 2002, Univeristy of Edinburgh, 8-10th July, 2002, pp. 23-29

[4] Langdon, P., Hwang, F., Keates, S., Clarkson, P. J., and Robinson, P. (2002), Investigating haptic assistive interfaces for motion-impaired users: Force-channels and competitive attractive-basins. In Proceedings of Eurohaptics 2002 (Edinburgh UK, 2002), pp. 122-127

[5] Roberts, J.C., Franklin, K., Cullinane, J. (2002), Virtual Haptic Exploratory Visualization of Line Graphs and Charts. In Mark T. Bolas, editor, The Engineering Reality of Virtual Reality 2002, volume 4660B of Electronic Imaging Conference, page 10. IS&T/SPIE, January 2002.

[6] Pokluda, L., Sochor, J., (2005) Spatial Orientation in Buildings Using Models with Haptic Feedback, Proceedings of WorldHaptics 2005 (Pisa, Italy, 2005), pp 523-524

[7] Wall, S., Brewster, S. (2004), Providing External Memory Aids in Haptic Visualisations for Blind Computer Users, In Proceedings of International Conference Series on Disability, Virtual Reality and Associated Tech-nologies (ICDVRAT), New College, Oxford, UK, 20-22 September 2004, pp. 157-164.

[8] Wood, J., Magennis, M., Arias, E.F.C., Helen Graupp, H., Gutierrez, T., Bergamasco, M. (2003), The Design and Evaluation of a Computer Game for the Blind in the GRAB Haptic Audio Virtual Environment, Euro-Haptics 2003, pp 147-158

[9] Magnusson, C., Rassmus – Gröhn, K. (2005), A Virtual Traffic Environment for People with Visual Impairment, Visual Impairment Research, Vol 7, nr 1, pp 1-12

[10] Prada, R., Payandeh, S. (2005), A Study on Design and Analysis of Virtual Fixtures for Cutting in Training Environments, Proceedings of WorldHaptics 2005 (Pisa, Italy, 2005), pp 375-380

[11] Saga, S., Kawakami, N., Tachi, S. (2005), Haptic Teaching using Opposite Force Presentation, Poster presen-tation (CD proceedings) WorldHaptics2005

[12] Magnusson, C., Rassmus-Gröhn, K.,Audio haptic tools for navigation in non visual environments ENACTIVE 2005, the 2nd International Conference on Enactive Interfaces, Genoa- Italy on November 17-18, 2005.

[13] Magnusson, C., Danielsson,H., Rassmus-Gröhn, K., Non Visual Haptic Audio Tools for Virtual Environments, Presented at the Workshop on Haptic and Audio Interaction Design University of Glasgow, 31st August - 1st September 2006.

Test of pan and zoom tools in visual and non-visual audio haptic environments

Charlotte Magnusson* Teresa Gutierrez*† Kirsten Rassmus-Gröhn*
(*)Lund University, Lund, Sweden
(†)LABEIN, Bilbao, Spain*
E-mail: charlotte.magnusson@certec.lth.se, tere@labein.es, kirre@certec.lth.se

Abstract

To enable visually impaired users to experience large virtual 3D models with a relatively small haptic device a test of a set of pan and zoom tools has been performed. The pan tools tested (keyboard, pressing the sides of the limiting box and click & drag using the PHANToM) were all seen to be useful. For the zoom tool the discrete keyboard press design was seen to work well, while the drag type zoom implemented worked poorly and will need to be redesigned. The test results show different preferences in visual and non-visual navigation, indicating the need for specially designed interaction utilities for the non-visual case.

1. Introduction

In most haptic virtual environments for the blind, the working area of the haptic device (such as the PHANToM) limits the size and complexity of the virtual environment that can be displayed. Haptic non-visual pan and zooming using a PHANToM device has been studied to some limited extent. Within the ENORASI study described in [1], a simple pan function was tested by a subset of the test users. Zoom functions to gain access to greater detail in virtual haptic line graphs are also suggested by Roberts et al [2], a test of a set of pan and zoom functions has been reported in [3] and this type of functions were also included in the virtual audio-haptic traffic environment described in [4]. The studies described in this paper were designed to further investigate pan and zoom utilities in a more realistic environment.

2. Test environment

The environment used was essentially the traffic model described in [4] (see figure 1), but with the model no longer divided into five parts, but instead joined into a single model.

Three different pan or scroll functions were implemented. They were similar to the ones used in [3]: moving the world using the arrow keys, moving the world by pushing the sides of the limiting box with the PHANToM stylus and moving the world by clicking & dragging using the PHANToM stylus.

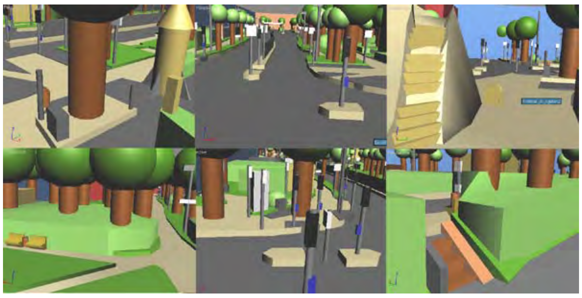


Figure 1. Views from the traffic environment model used.

When moved by pressing the arrow keys, the world moved in the direction of the key pressed. When the user pushed the limiting box, the world moved as if the box was actually pushed by the user over the model. For the click and drag pan function the world became attached to the PHANToM stylus, although to avoid unpredicted movements in the vertical direction the movement was restricted to the horizontal plane. The two discrete move functions (arrow keys and limiting box) allowed for different move lengths. The move length that should be used was selected by the user.

For the zoom we used two basically different implementations: stepwise zoom by pressing a key on the keyboard and click & drag zoom using the PHANToM stylus. The click & drag zoom was implemented based on a rubber band metaphor i.e two points can be dragged with respect to each other to enlarge or reduce the size of the model (like you would to enlarge or reduce the size of a rubber band). Since the interaction device (the PHANToM) only has one interaction point, the position of the first point was indicated by a point and click action, and then the user pointed the stylus at the second point, held down the switch and used drag to resize the world.

The stepwise zoom used fixed zoom factors of 0.8 for zooming out and 1.25 for zooming in. It was implemented so that the tip of the PHANToM stylus was always in the same position relative to the world before & after the zoom. To test the importance of knowing the contact point with respect to the world, users were asked to do the stepwise zoom in two different ways: in contact with the model and lifting the stylus a distance above the world.

3. Test setup

The test was performed by 12 sighted users (age 23-60). To speed up the learning process and also to compare sighted interaction with interaction using only touch and hearing the test users first performed the test with vision and after that without visual feedback.

The order in which the tools within the pan and zoom groups were tested was permuted to avoid learning effects. The pan tools were tested before the zoom tools both in the visual and the non-visual case.

All test persons started the test with visual feedback, and were initially allowed to familiarize themselves with the environment and with the different move lengths. The move lengths were not tested individually, but instead the users were instructed to use whatever lengths they wanted and the usage of different length settings was recorded. After the initial familiarization, the pan test task was performed for all three pan tools. The pan task was to locate the bus stop in the south part of the environment and then to move the PHANToM stylus to a rocket in the north part of the environment while making use of the specified pan tool when panning was needed. The initial position of the world was always the same (showing the middle part of the world).

The zoom test task was to first locate the bus stop and then to zoom in and point to 3 objects at the bus stop: the waste paper basket, the sofa and the stairs. This task was done for all three zoom tools.

As a final task the user was asked to do the pan and the zoom tasks once again but this time using his or her favorite tools – the user could use any of the tested tools, and the tool use was recorded.

After this, all these test tasks were performed again but without visual feedback.

4. Results

All users were able to complete the test tasks both with and without visual feedback. The results were analyzed with analysis of variance (ANOVA) where the independent within-group variable was the navigational strategy used.

Post hoc tests were done using the Tukey test and the significance level was set to 0.05 throughout the analyses.

Pan method	Mean (s)	Sd (s)
Arrow keys, vision	112	82
Limiting box, vision	111	50
PHANToM drag, vision	92	46
Arrow keys, no vision	447	346
Limiting box, no vision	320	109
PHANToM drag, no vision	372	234
Favorite pan, vision	76	38
Favorite pan, no vision	198	91

Table 1. Means and standard deviations (Sd) for the time to complete for the different pan tasks.

For the pan tools, the ANOVA for the dependent variable time to complete (see table 1) revealed significant differences ($F(7,77)=14{,}1$, $p<0.05$). As expected, the post hoc test showed significant differences between all three pan methods with vision and all the three non-visual pan methods, while the difference between the favorite pan with and without vision, although showing a tendency in this direction, did not turn out to be significant ($Q(8,77)=3{,}2$). In the non-visual condition the use of the arrow keys and the PHANToM drag was significantly slower than the final test where the favorite tool(s) was(were) used ($Q(8,77)=6.5$, 4.5). A tendency ($Q(8,77)=3{,}2$) was seen also for the comparison with the limiting box. No such effect was seen for the visual case which indicates a stronger training effect for the non-visual interaction. As expected all the non-visual pan methods were all significantly slower than the favorite pan with vision.

In the visual condition, the favorite tool selected for the final task by 8 users was the arrow keys while 5 users preferred the drag (one user used both). None of the users preferred pushing the limiting box for interaction in the visual case.

For the non-visual condition, the tools preferred by the users in the final pan test task are different. For non-visual interaction 6 users preferred the PHANToM drag, 4 users preferred the arrow keys and 2 users preferred pushing the limiting box.

For the discrete pan tools, the most often used move factors were 10, 20 and 60 in both conditions. The results in the non-visual condition were a bit more pronounced (probably due to the fact that the users now were more familiar with the environment) but the same move factors were most frequently used. The actual length of the move was calculated as 0.0016 m (the proxy radius) times the move factor. In the current environment a move factor of 60 corresponded to a page up/page down type of move: if the environment was moved in the north – south direction, the move was done so that the points visible at a bottom south position would be visible at a top north position after the move.

Zoom method	Mean (s)	Sd (s)
Keyboard no contact, vision	94	34
Keyboard with contact, vision	87	62
Zoom drag, vision	135	102
Keyboard no contact, no vision	196	64
Keyboard with vision, no contact	168	63
Zoom drag, no vision	215	105
Favorite zoom, vision	66	32
Favorite zoom, no vision	144	91

Table 2. Means and standard deviations (Sd) for the time to complete for the different zoom tasks.

The results for the zoom are summarized in table 2. The zoom drag turned out to be hard to use, which also shows up in the post hoc test. While both keyboard zooms were significantly faster than all the zoom tools without vision, this was not the case for the zoom drag. The tendency for the favorite zoom with vision to be faster than the non visual favorite is more pronounced than for the pan tools ($Q(8,77)=4.2$) but it is not significant (the significance level is at 4.4). Comparing the zoom tools with the corresponding favorites did not show any significant differences (less of a training effect). This could be due to the fact that the zoom tools were tested after the pan tools – but the task for the zoom operations also appeared to be less favored by the visual interaction (some of the objects the user should locate were usually hidden from view). In the non-visual condition, one very experienced PHANToM user liked the zoom drag tool (due both to the fact that this tool provided some feedback about the amount of zoom and to the fact that one only needed to zoom once).

Finally, if we look only at the mean values, we see that after some training and with a choice of interaction techniques the completion times for the non-visual case are only 2-3 times longer than the times obtained with vision.

5. A related study

A related study was performed at LABEIN, Bilbao, Spain.

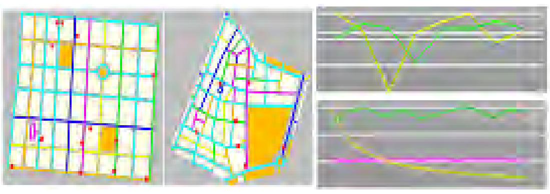


Figure 2. The types of graphs used in the LABEIN test.

This test was done using the GRAB [5] two-point haptic device, and tested the way visually impaired users were able to interact with audio-haptic maps and graphs (figure 2). The applications used in this test include pan and zoom utilities, it provides an interesting complement to the results presented above.

5.1. Pan and zoom implementation

The pan function in the LABEIN test was implemented so that the position of the control finger was fixed relative to the virtual graph scene. The movement of the control finger was then used for translating the scene. While the users were panning, they received a resistance force to indicate that the workspace was being moved. The user was unable to move outside the graph scene. This implementation was similar to the click & drag panning tested in the Lund test (although in Lund you could pan outside the model). The zoom function implemented for the LABEIN tests was designed in a similar way as the stepwise zoom function used in the Lund test. As the size was changed by pressing a keyboard key the control finger remained in the same position relative to the world, while the other finger would end up in a different position due to the size change.

5.2. Test results

The LABEIN test was performed with visually impaired users – the map application was tested by 4 congenitally blind persons and 5 persons with low vision, while the graphs application was tested by 4 congenitally blind persons and 6 persons with low vision. The users were allowed a training session before the actual test to familiarize themselves with the applications. For the graphs application 7 users thought the advanced utilities such as pan and zoom were easy to use while 3 thought they were "not too easy but not too hard". For the maps application 8 users thought these utilities were easy to use and one user thought it was "not too easy but not too hard".

Looking at the user feedback received, two users with low vision stated that the panning was not intuitive. These users were used to work with magnifiers so they hoped to move the workspace (or the tool) and not the graph. Some users found it hard to realize that a graph did not fit in the workspace (indicating the usefulness of feedback on this point). It was also seen that it could be useful to be able to select the range of the zooming and scaling, or at least to provide the user with feedback about the range.

6. Discussion and conclusion

The results of the Lund traffic environment pan and zoom test show that all the pan function designs are working quite well. In contrast with [3], it now turns out the PHANToM drag is the most popular pan tool for non-visual interaction. The usefulness of the drag function is further strengthened by the results obtained at LABEIN. Still, all three tool designs were shown to be useful (as in [3]). One problem with the arrow key panning is to know which way the world will move. In the present implementation it is the model that moves in the direction of the arrows. Some users initially had problems with this, which indicates the importance of providing the user with feedback of what has happened. In Lund no users had problems with the direction of the click and drag type of panning, while in the LABEIN test two users had problems with the panning direction for this type of panning. The test result shows the need of at least two (maybe three) different move lengths for the discrete pan operations. One really large, like a page down/page up and one shorter for the finer adjustments is needed. In the current test the proportions for the three most used lengths were 1:2:6.

One problem was the fact that it was possible to move the model far away from the work space. This indicates that when the size of the world is known an environment like this should simply not allow the model to move completely out of the workspace (already implemented in the LABEIN applications).

For the zooming operations it was clear that the design where the user is in contact with the same point of the model before and after the zoom works well. Despite this, several users complained about the lack of feedback on how much the world was zoomed, and more feedback on this point needs to be added. Both these results agree with the results from the LABEIN test. The design of the drag zoom function was not a success. Despite this the user responses indicate that some kind of direct manipulation type zoom could be quite useful. Further investigation is clearly needed on this point.

The results obtained also support the earlier study by Jansson et al [6] which show the importance of training – our results for the pan tools indicate a training effect particularly for the non-visual interaction. The Lund test highlights the fact that user preferences depend a lot on whether or not they have access to visual feedback. This is a reminder for anyone developing non-visual applications that the interaction needs to be specially designed for the non-visual case.

Finally this test confirms that it is quite possible to understand and interact with large and complex haptic environment also in the non-visual case.

7. Acknowledgements

The study was carried out with the financial assistance of the European Commission which co-funded the IP “ENABLED”, FP6 - 2003 - IST - 2, No 004778 and the NoE ENACTIVE - “ENACTIVE Interfaces”, FP6-2002-IST-1, No IST-2002-002114. The authors are grateful to the EU for the support given in carrying out these activities.

References

[1] C. Magnusson, K. Rassmus-Gröhn, C. Sjöström and H. Danielsson, “Navigation and recog-nition in complex haptic virtual environments – reports from an extensive study with blind users,” Eurohaptics 2002, Edinburgh, UK

[2] J.C. Roberts, K. Franklin, and J. Cullinane, “Virtual Haptic Exploratory Visualization of Line Graphs and Charts”, The Engineering Reality of Virtual Reality 2002, Mark T. Bolas, editor, volume 4660B of Electronic Imaging Conference, page 10. IS&T/SPIE, January 2002

[3] C. Magnusson, K. Rassmus-Gröhn, “Non-visual Zoom and Scrolling Operations in a Virtual Haptic Environment”, EuroHaptics 2003, Dublin, Ireland

[4] C. Magnusson, K. Rassmus-Gröhn, “A Virtual Traffic Environment for People with Visual Impairments”, Visual Impairment Research, pp 1-12 Vol 7. No 1. 2005

[5] R. Iglesias, S. Casado, T. Gutiérrez, J.I. Barbero, C.A. Avizzano, S. Marcheschi, M. Bergamasco, “Computer graphics access for blind people through a haptic and audiovirtual environment”, HAVE 2004 (Ottawa, Canada, 2-3 October 2004)

[6] G.Jansson and A. Ivås, Can the efficiency of a haptic display be increased by short-time practice in exploration? Proceedings of Haptic Human-Computer Interaction 2001

Perceiving what is within reach from intermodal patterns

Bruno Mantel[*] Benoît G. Bardy[*] Thomas A. Stoffregen[†]
(*)Motor Efficiency & Deficiency Lab, Montpellier 1 University, France
(†)Affordance Perception Action Lab, University of Minnesota, U.S.A.
E-mail: bruno.mantel@univ-montp1.fr, benoit.bardy@univ-montp1.fr, tas@umn.edu

Abstract

Studies on distance perception often focus on visual cues, and to a smaller extent on haptic or auditory cues. Here we argue that perceptual information about egocentric distance exists not solely within modalities but also extends across modalities (i.e., in intermodal patterns). We formalized an intermodal invariant relation specifying the distance to a static object for a moving observer. In a series of three experiments, we used a virtual reality set-up to delete all traditional visual cues about distance and tested whether scaled information about egocentric distance could be perceived from our intermodal invariant relation. Experimental results confirm our hypotheses and call for a better consideration of intermodal patterns in perception studies as well as for the design of interfaces.

1. Introduction

Behavior causes simultaneous changes in multiple forms of ambient energy (e.g., light, sound, pressure, etc.). Research on multimodal perception has tended to focus on the integration (in the nervous system) of stimuli available to different sensory systems. At the same time, research on layout perception has investigated several modal cues to distance (visual, auditory, haptic, etc.). However, potential sensory stimulation is not limited to patterns available to individual sensory systems. There are also patterns extending *across* different forms of ambient energy which provide relational information that is not present in their constitutive parts. These superordinate, higher-order patterns make up the *global array* [4, 5].

We decided to evaluate perceptual sensitivity to the global array in the context of egocentric distance perception. We formalized a relation across optics and inertia providing information about egocentric distance to a moving perceiver. The apparent angular motion of the target did not provide, on its own, any scaled information about ego-distance [1, 2, 3]. On the other hand, the invariant relation across optical and inertial energies was related to egocentric distance: in short, the optical apparent displacement of the target was scaled in terms of the inertial consequences of head movement.

We conducted three experiments to test whether that intermodal relation can be used by humans to detect whether objects are within reach.

2. General method

2.1. Task

Participants, seating in a dark room, were presented static virtual targets at different distances. Their non-dominant eye was patched and their field of view was reduced to approximately 50°. Their task was to judge verbally ("yes"/"no") whether the presented virtual object was reachable or not ("if you were extending your arm, without leaning forward, will you touch the target with your index finger?"). In the 2nd and 3rd experiment, participants were also asked to evaluate (between 1 and 5) how confident they were about their judgments.

2.2. Apparatus

The target was a solid flat cylinder (coin-like) which angular size was kept constant across trials. It was displayed at eye-height on a large video projection screen, located 1.20m away in front of the participant (experiment 1 and 2) or on the LCD matrix of a Head Mounted Display (HMD, experiment 3). In both set-ups, displays of the target were driven in real time by the movement of the participant, using our intermodal invariant relation. This allowed us to simulate the virtual target at different distances, as illustrated in the top schema of figure 1, while controlling for other possible cues. For that purpose and for analyses a posteriori, participants' movements were sampled at 100 Hz with a 6 degrees of freedom electromagnetic sensor, providing a linear resolution smaller than 1mm, and an angular resolution smaller than 1°.

2.3. Design

2.3.1. Preliminary session

Several anthropometric measurements were taken prior to the experiment and used to build an anthropometric model for estimating each participant's reaching capabilities. We also measured each participant's actual maximal reachable distance (or critical boundary, CB) using the method of limits: participants were moved backward of forward from the screen, by increments of 1 cm (2 ascending series, 2 descending series), and had to reach from each location for a flat object. The measure of the actual maximal reachable distance was then used as a reference for building the body-scaled set of distances at which the target had to be judged.

2.3.1. Verbal judgments session

The number of target distances, the number of trials at each distance, as well as the range of distances used are detailed in Table 1. In the second and third experiments we had 2 groups of subject which differed by the range of distances they had to judge.

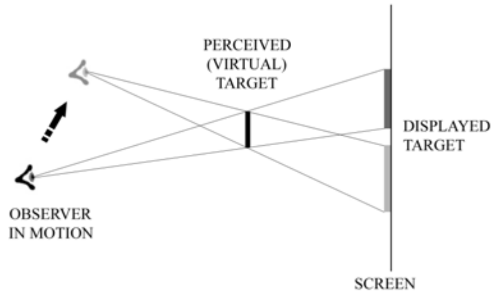
Table 1. Experimental design

	Exp.1	Exp.2	Exp.3
Number of distances	5	13	17
Distance range (%)*	90-110	44-128 72-156	23-135 72-184
Trials at each distance	4	6	5

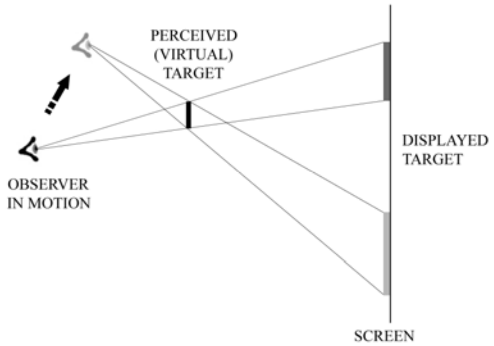
* Distances are body-scaled and expressed in percentage of actual maximal reachable distance.

In the first experiment, we tested 6 different experimental conditions. Figure 1 presents an overview of 5 of these conditions.

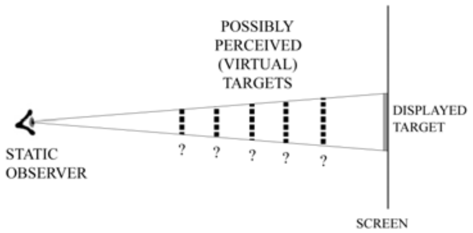
In the reference condition, Vision-Movement (VM), participants were allowed to explore the scene freely (by moving their head and/or torso) prior to giving their judgments. Hence, in that condition, all terms of the intermodal invariant specifying ego-distance were present and we expected participants to perform the task successfully.



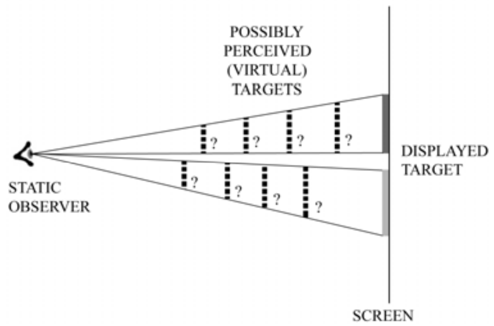
Vision-**M**ovement condition (VM)



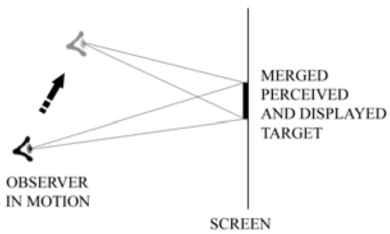
Vision-**M**ovement with **g**ain condition (VMg)



Vision-**O**nly condition (VO)



Vision-**O**nly-**p**lay**b**ack condition (VOpb)



real-**V**ision-**M**ovement condition (rVM)

Figure 1. Experimental conditions.

In the Vision-Only (VO) condition, participants had to remain static and therefore should not have any idea about how far the target was. We also used a gain condition (VMg) in which we introduced a gain in the relation used to drive display according to movements. We expected perceived distance to be shifted according to the gain value.

In the playback condition (VOpb), participants were static but their movements were played back from earlier trials and used to drive the display of the target. As we have seen, passive motion parallax should not be sufficient on its own to perceive distance. At last, we also had two conditions with real vision instead of virtual (with movement, rVM, and without movement, rVO), to validate our virtual set-up. In these conditions the distance between the target and the participant was not varied virtually but actually, by physically moving the participants between trials.

2.4. Dependent variables

Psychometric curves were fitted to yes/no responses, using maximum likelihood estimates as deviance measure, Monte Carlo simulations to assess the goodness of the fit and bootstrap method for calculating confidence intervals [6, 7]. From the fitted curves we derived two main parameters: the slope (taken at Y=50% of "yes" responses), characterizing the sharpness of the transition and hence the *precision* of judgments; and the so-called threshold (abscissa of the point at Y=50%), indicating the locus of the transition and therefore representing the *accuracy* of judgments. We also analyzed the confidence ratings.

4. Results and discussion

In all three experiments, participants perceived egocentric distance accurately only when the intermodal relation between optics and inertia was preserved. The slopes obtained in VM condition were each time significantly different from those obtained in VO and VOpb conditions. Indeed, the VO slopes were almost zero, indicating an horizontal curve with no transition, while the slopes of the playback curves were always at least four times smaller (and far more variable) than the intermodal (VM) slope. Accordingly, when static subjects were shown play-back optics as open-loop stimuli, judgments of ego-distance were grossly inaccurate, confirming that optical motion parallax was not sufficient.

The slopes obtained in the two other intermodal condition, namely Vision-Movement-with-gain (VMg) and real-Vision-Movement (rVM), were very similar (and not significantly different) from those obtained in the reference condition (VM) but were significantly different from, and far greater than, VO and VOpb slopes.

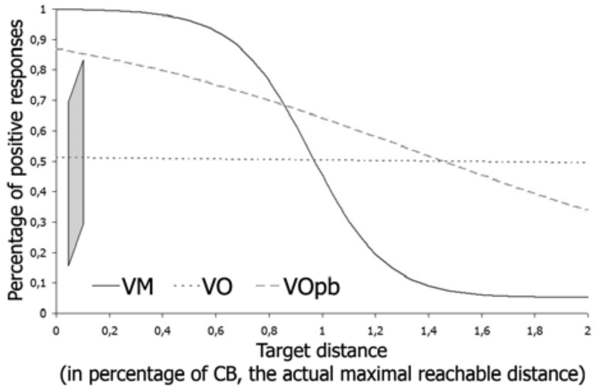


Figure 2. Regression curves fitted to judgment data. The percentage of "yes" responses is expressed as a function of target distance (body-scaled, see method section). The grey rhombus on the left represents the actual location of the target's display at all times, its distance from the observer being only varied virtually (*intermodally*).

Concerning the accuracy, as predicted, changing the gain in the intermodal relation shifted the perceived distance accordingly. The ratio between perceived distance in intermodal conditions with and without gain was close to the value of the gain. Similarly, the reference VM curve and the curve from real Vision-Movement were overlapping almost perfectly, thus validating our virtual set-up.

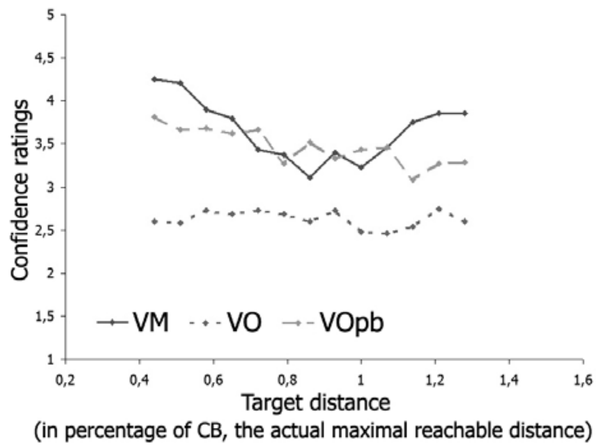


Figure 3. Regression curves fitted to judgment data. The percentage of "yes" responses is expressed as a function of distance (body-scaled, see method section).

Analyses of variance performed on confidence ratings revealed significant main effects of condition and distance as well as a significant interaction for three of the four groups of participants (2 experiments x 2 groups with different distance range), while for the 2nd group of experiment 3, only the main effect of condition was significant. In all four groups the curves had similar shapes. The confidence curves corresponding to the Vision-Only condition were almost linear and horizontal, while for Vision-Only-playback condition the curves were still linear, but slightly decreasing, indicating a small progressive loss in confidence as distance increased. Conversely the curves for Vision-Movement (intermodal) condition were all U-shaped indicating a higher confidence when targets were very far or very close and a lower confidence for targets in between, where the situation is more ambiguous. Figure 3 illustrates these curves with the example of group 1 from experiment 3.

In conclusion, there is intermodal information, across optics and inertia, specifying whether an object is reachable or not, and that information is not present at all in individual energies. Human participants are sensitive to that intermodal invariant relation and, on this basis, detect accurately egocentric distance in terms of its reachability.

References

[1] Bingham, G. P. & Stassen, M. G. (1994) Monocular egocentric distance information generated by head movement. Ecological Psychology, 6, 219-238.

[2] Eriksson, E.S. (1974). Movement parallax during locomotion. Perception and Psychophysics, 16, 197-200.

[3] Sedgwick, H. (1986). Space perception. In K.R. Boff, L. Kaufman, & J.P. Thomas (eds.) Handbook of Perception and Human Performance, Vol 1, pp. 21:1-57. New York: Wiley.

[4] Stoffregen, T.A. & Bardy, B.G. (2001). On specification and the senses. Behavioral and Brain Sciences, 24(2), 195-261.

[5] Stoffregen, T.A. & Bardy, B.G. (2004). Theory testing and the global array. Behavioral and Brain Sciences, 27(6), 892-900.

[6] Wichmann, F. A. & Hill, N. J. (2001): The psychometric function: I. Fitting, sampling and goodness-of-fit. Perception and Psychophysics, 63(8), 1293-1313.

[7] Wichmann, F. A. & Hill, N. J. (2001): The psychometric function: II. Bootstrap-based confidence intervals and sampling. Perception and Psychophysics, 63(8), 1313-1329.

Acknowledgements

Research supported by Enactive Interfaces (European IST contract #002114).

Can Humans Learn How to Minimize Unintended Interpersonal Coordination?

Marin L.[*] Issartel J.[*†] Cadopi M. [*]
(*) Motor Efficiency and Motor Deficiency Lab, University Montpellier 1, Montpellier France
(†)Sonic Arts Research Center, Queen's University, Belfast, Northern Ireland
E-mail: ludovic.marin@univ-montp1.fr, j.issartel@qub.ac.uk, marielle.cadopi@univ-montp1.fr

Abstract

When two persons have perceptual contact their movements are unintentionally coordinated (called unintentional interpersonal coordination). This unintended synchronization seems to be a natural process of any dyadic behavior between two humans. This study aims to investigate whether humans can learn how to minimize this unintended phenomenon of interpersonal coordination.

Expert dancers and non-dancers were instructed to freely move their arm in two different conditions. In the "Alone" condition each participant was alone whereas in the "Paired" condition two participants facing each other were instructed to not take into account the movements of the other participant when performing his/her own moves.

The main result showed that although dancers were able to not coordinate with the other participant in the Paired condition, non-dancers could not. These findings reveal that contrary to non-dancers, dancers have learned a new way of dealing with the spontaneous process of interpersonal coordination.

1. Introduction

Two persons acting together generate interpersonal coordination. This phenomenon is present in our everyday life as soon as two people have perceptual contact (visual, tactile…) [5]. Even if interpersonal coordination can be intentional (soldiers marching at the same pace), in our daily life it is mostly unintentional. For instance, two persons walking together in the street synchronize their step frequency even though they are not aware of the synchronization. Similarly, although people have their own hand clapping frequency, everyone tends to applause at the same (synchronized) pace when one person starts clapping [4].

Experimental evidence of unintentional coordination was brought to light by Schmidt and O'Brien in 1997. In this study the authors asked participants to swing a pendulum at their preferred frequency in two different conditions: with or without having the other participant's moves in sight. The results showed that when participants saw the other participant's movement, they both immediately coordinate together. The authors concluded that unintended coordination emerges even if the synchronization was not the goal of the situation (or the instructions). In the line with the previous results, Issartel, Marin and Cadopi [2] conducted an experiment where pairs of participants were instructed to intentionally *not* coordinate their movements with the other participant. Their findings indicated that participants were unable to not coordinate with each other. Together these experiments show that the emergent phenomenon of coordination is so powerful that humans cannot avoid an unintended dyadic motor co-ordination. Consequently, this phenomenon is an illustration of the spontaneous process that underlies interpersonal coordination.

But are humans able to change their natural process? Can we learn how to minimize the unintentional phenomenon of interpersonal coordination? In this study we are interested in investigating, if humans ***can*** adapt new "ways" of dealing with this spontaneous process. We hypothesize that expert dancers can be a representative example of humans that are able to adapt new "ways" of dealing with this spontaneous process. We predict that expert contemporary dancers are able to *not* coordinate with someone if we ask them to do so because during long hours of rehearsal they are used to acting without taking into account the moves of other dancers.

2. Method

Twelve pairs of participants (6 pairs of expert contemporary dancers and 6 pairs of non dancers) sat on a chair with their right elbow resting on the surface of the table while the left arm rested on their lap. They were instructed to freely move their right forearm in the sagittal plane without constraints about frequency or amplitude moves.

Such free movements are characteristics of what dancers call an improvisational situation. Expert contemporary dancers are often involved in an improvisational dance situation, which is a very rich situation from a dynamical point of view: 1) movements are not planned (movements emerge), 2) movements are made from moment to moment and 3) improvisation is made straight away. In general an improvisation task reveals dancer's individual properties (referred to as a motor signature, [2]), which are manifested as a whole range of preferred frequencies. Besides individual properties, when two or more dancers are involved in an improvisational choreography, a collective signature emerges from the interaction between dancers.

Participants performed two different improvisational conditions. In the first condition (called *Paired*), instructions were explicitly to *not coordinate* his/her movements with those of the other participant. The participant had to move as if he/she were alone, without taking into account the movements of the other to perform his/her own movements. In the second condition (named *Alone*) the participant was alone in order to assess the individual motor signature without any influence of the other participant. Each participant performed one block of 6 trials for each condition (conditions were counter-balanced). The duration of each trial was of 3 minutes with a 2-minute rest interval between trials.

Analyses were performed on the time variation of the elbow angle at a sampling rate of 50 Hz. Data was collected with elbow electrogoniometers Biometrics SG 110 (Biometrics, Oxford, England). Investigating improvisational movements is a challenge because it leads to non-stationary data: extreme changes of frequency and amplitude in the signals, and multiple frequencies produced at the same time. Classical methods cannot be used to analyze such data. Consequently, we applied methods that take into consideration these signal property changes. Wavelet Transform (WT) and Cross-Wavelet Transform (CWT) methods – applied to the conjoint analysis of two signals - are well suited to study more natural time-series since the WT and the CWT methods are able to take into account multiple frequencies and non-stationary signals. These methods are required to study multi-scale signals occurring over finite spatial and temporal domains. Hence, the WT and CWT methods allow us to analyze the whole range of frequencies and the temporal evolution of each frequency of non-stationary signals. To identify the significant structures from a noisy component of the WT and CWT analyses, Torrence and Compo [6] developed a method of defining a confidence level (grey lines in Figure 1) above which a maximum in the local WT (and CWT) spectrum is statistically significant (desired level in this article: 95% confidence). The data used to quantify the movements are extracted from these significant areas. A program written in the software language C^{++} transformed raw data into WT and CWT. The analysis of the signals was performed with the Morlet complex mother function (order of 8, see [1]). To cover the frequency range of interest, a large band of frequencies were chosen (0.04 to 6.25 Hz) for the whole analysis. Data were analyzed using Matlab software (MATLAB 7.0, the Math-Works Inc., Natick, MA).

Once the data were extracted from the WT and the CWT analyses, we performed two different investigations: one on the motor signature of each participant and one on interpersonal interaction within each pair.

The motor signature investigation was performed on the WT analysis. We assessed the whole range of frequencies that each participant performed in each trial. The whole range of frequencies provides the individual properties of a participant in one trial. If the same individual properties are consistent in every trial, we can claim that the given range of frequencies embodies the motor signature of the participant. In such a context, we tested the reproducibility of two variables. 1) For each trial, we performed for each frequency the sum of the WT spectrum. From this, we obtain a *distribution of the frequency* spectrum (D-WT) that represents the intensity of each frequency in one trial. 2) We also measured the *number of frequency occurrences* (NF) of the WT spectrum for one trial. The comparison of the D-WT and the NF of each trial by condition allowed us to estimate the reproducibility of these two variables and consequently, of the *motor signature*. For a complete motor signature analysis, this reproducibility was calculated with a cross-correlation analysis between each trial for each condition and also between conditions.

The interpersonal interaction within each pair of participants was analyzed with the CWT, which represents the common frequencies between two signals. To identify whether any frequency interaction occurred, we analyzed the variable inter-condition D-CWT, similar to the previous variable D-WT. D-CWT compares the *distribution of the frequency* spectrum between two participants.

In order to compare the differences between the two groups, we performed non-parametric analyses (Friedman's test and Wilcoxon Signed Rank tests) on all variables presented above.

3. Results

3.1. Visual analysis of the motor signature

A visual analysis of participants' *motor signature* from WT spectrums is presented in Figure 1. This figure is an illustration of three typical trials (trials 1, 3

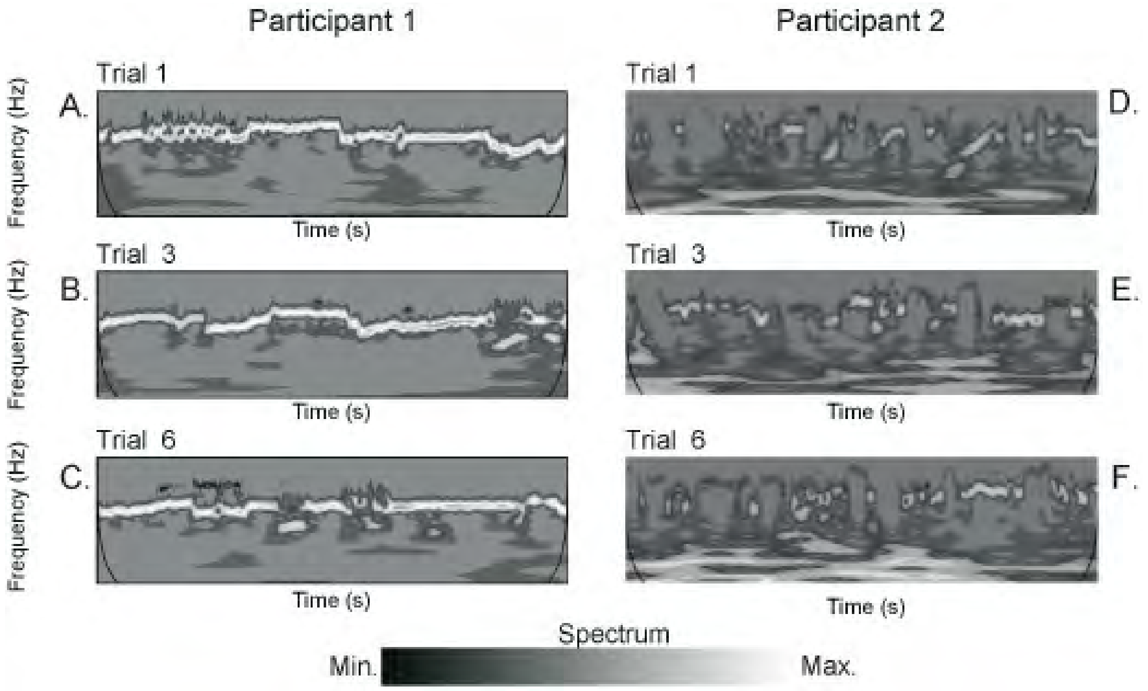


Figure 1. Visual analysis of two typical participants' motor signature. Grey lines represent significant frequencies. We can see that each participant maintains his/her motor signature regardless the trial performed

and 6 of *Paired* condition) for two participants (participant 1 on the left column and participant 2 on the right). As we can observe, each participant maintains the same frequency range. Whatever the trials, the grey lines (significant areas) are quite similar. The visual inspection reveals that participants have a consistent motor signature: each participant maintains the same frequency distribution of movements and the same number of frequency occurrences.

3.2. Quantitative analysis of the motor signature

The visual analysis is confirmed by quantitative analyses. Friedman's tests on the D-WT inter-trial within each condition reveal that there are no statistical differences for both groups (For dancers, *Alone*: X^2 = 17.68, df = 14, $p > .05$; *Paired*: X^2 = 06.57, df = 14, $p > .05$; For non-dancers, *Alone*: X^2 = 21.29, df = 14, $p > .05$; *Paired*: X^2 = 13.97, df = 14, $p > .05$), indicating a similar frequency distribution for all trials in each condition. More interestingly, there is no statistical difference on the D-WT between the two conditions for both groups: For dancers $z = -0.31$, $p > .05$; For non-dancers $z = -0.23$, $p > .05$. In line with D-WT inter-trial and inter-condition results, Wilcoxon Signed Rank tests on NF show that there is a similar number of frequency performed by participants in inter-conditions for both groups: For dancers, $z = -1.41$, $p > .05$; For non-dancers, $z = -0.94$, $p > .05$. Together these results indicate the presence of a motor signature for each participant as trials have a similar distribution and number of frequencies within each condition. The cross-correlation analysis of each trial in all conditions confirms once again that all dancers and non-dancers maintain their motor signature (R = 0,9 dancers; R = 0,73 non-dancers). These findings illustrate that even if participants can freely move without frequency and amplitude constraints, dancers' and non-dancers' movements are limited to a preferential range of frequencies.

3.3. Analysis of interpersonal interaction

In order to analyze interpersonal interaction, for each pair, we perform cross-correlations on D-CWT (which compares the *distribution of the frequency* spectrum within a pair). The results on the non-dancers group show a significant difference between Alone and Paired conditions ($z = -2.20$, $p < .05$). The mean of the Alone condition for non-dancers reveals low correlation coefficients of R = 0.37. Each participant was alone when he/she performed this condition. These low correlation coefficients are explained by a random production of the same frequencies. For Paired condition, results show a mean of correlation coefficients of R = 0.66. Although participants were instructed to intentionally *not* take into consideration movements of the other, an unintentional synchrony emerged. For the dancers group, results are interestingly different from the non-dancers group: there are no statistical differences between Alone ($R = 0.86$) and Paired ($R = 0.82$) conditions. These findings reveal that contrary to non-dancers, there is no emergence of unintentional interaction for expert dancers.

4. Discussion

In this study we have found two main results. First, there is a strong hold of the motor signature for dancers and non-dancers. Interacting with someone does not change the individual motor signature for dancers and non-dancers. Note, however, that for non-dancers interaction modulates the variety of some frequencies. These modulations illustrate the influence of the Paired condition on the motor signature. Nevertheless, this influence does not radically modify the intrinsic individual property. Regardless the level of dance expertise, this *motor signature* can be understood as the expression of the intrinsic dynamic [3] that leads the motor behavior in a specific and limited range of frequencies.

The second finding of this study and the most significant is that dancers are able to intentionally not be coordinated with the other dancer whereas the non-dancers cannot. During long hours of rehearsal expert dancers have learned to perform their own specific choreography *without* taking into account the environment around them *(i.e., other dancers, spectators)*. Conversely, although instructions explicitly asked non-dancers to *not* coordinate their movements with those of the other participant, non-expert dancers are unable to follow these instructions. The perceptual contact is so strong that unintentional coupling occurs between the two individual movements.

Even if all humans seem to use the same processes to synchronize (e.g., non-dancers in this study or participants in Schmidt & O'Brien's study), the example of expert dancers shows that some humans are able to not follow the same processes as non-expert dancers - they have found new ways of dealing with this synchronization. To answer our previous questions: are humans able to change their natural process? Can we learn how to minimize the unintentional phenomenon of interpersonal coordination? The answer is definitively yes. This study is an illustration of the wonderful capacity humans have for adaptation. This level of sophistication and complexity is unique to the human species, and we have yet to see these kinds of developments in robots or machines

5. References

[1] Issartel, J., Marin, L., Gaillot, P., Bardainne, T., and Cadopi, M. "A Practical Guide On Time-frequency Analysis To Study Human Motor Behavior: The Contribution Of Wavelet Transform", Journal of Motor Behavior, pp. 139-159, 38, 2006.

[2] Issartel, J., Marin, L., and Cadopi, M. "Unintended Interpersonal Co-ordination: Can We March To The Beat Of Our Own Drum?", Neuroscience Letters, pp. 174-179, 411, 2007.

[3] Kelso, J.A.S. Dynamic patterns: The self-organization of brain and behavior, MIT Press, Cambridge, MA, 1995.

[4] Neda, Z., Ravasz, E., Brechet, Y., Vicsek, T., and Barabasi, A.-L. "The Sound Of Many Hands Clapping", Nature, pp. 849-850, 403, 2000.

[5] Schmidt, R.C., and O'Brien, B. "Evaluating The Dynamics Of Unintended Interpersonal Coordination", Ecological Psychology, pp. 189-206, 9, 1997.

[6] Torrence, C., and Compo, G.P. "A Practical Guide To Wavelet Analysis", Bulletin of the American Meteorological Society, pp. 61-78, 79, 1998.

A Powerful Low-Cost Visualization Setup

Giuseppe Marino   Davide Vercelli   Franco Tecchia
Paolo Simone Gasparello   Marcello Carrozzino   Massimo Bergamasco

PERCRO Lab, Scuola Superiore S. Anna, Italy
E-mail: giuseppe.marino@sssup.it

Abstract

We introduce a low-cost setup for immersive and interactive visualization of complex virtual environments.

A high resolution stereoscopic image is obtained by combining the images coming from four SXGA+ projectors onto a screen structure of our design. Overlapping areas and distortions are handled in software by our system. A commercial tracker is used both for the user's point of view and a 6DOF wand for further interaction. Applications built on top of the XVR authoring environment benefit from a network rendering software that spreads the rendering task transparently among a cluster of commodity computers, without requiring the original application to be aware of it.

We also describe an example application to show how this setup may be used as an aid in an industrial design setting.

1. Introduction

It is commonly accepted that Virtual reality can be defined as interactive computer graphics that provides viewer-centred perspective, large field of view and stereo. While Head-mounted displays (HMDs) achieve these features with small display screens, placed close to the viewer eyes, which move with the viewer, projection-based displays, supply these characteristics by placing large, fixed screens more distant from the viewer. When a large and complex 3D model is projected on a large surface, projection system may suffer for the limited resolution offered by the single projector, unless dedicated and expensive beamers are used. A way around this limitation is to use multiple projectors to compose a tiled image: in this way there is virtually no limitation to the resolution of the image presented to the user, but typically requires a dedicated rendering architecture that takes care of synchronising a cluster of PC connected to multiple projectors. A number of solutions to obtain low-cost interactive projection-based displays was presented in literature [2][7].

In the following we describe how we accomplished this task in a system based on the XVR technology, a VR software framework commonly used in the context of the ENACTIVE NoE. The resulting system constitute a cost-effective solution to the presentation of high-resolution and highly interactive 3D graphics content, providing a simple-to-use general architecture able to accommodate even very complex VR applications.

2. XVR

XVR [1] is a stand-alone Integrated Development Environment (IDE) for the rapid development of complex Virtual Reality applications.

Development started in 2001 and today XVR is used for a wide range of projects dealing with interactive graphics, interaction and physics simulation. As a result, the XVR framework gets continuously updated to accommodate the evolving needs of the user community.

Besides real-time graphics, XVR offers a wide range of features to control many collateral aspects of VR programming, including sound, interaction and haptics.

XVR applications are developed using S3D (Script 3D), a dedicated scripting language with a Javascript-like syntax; via a single, unified programming language it is possible to deal with all the various aspects of the system, from 3D animation to haptic interaction.

3. XVR Network Renderer

XVR Network Renderer [5], takes advantage of a network of calculators to perform rendering, according to the "sort-first" cluster rendering approach [6].

Each machine takes care of a subset of the rendering task, thus allowing large output resolution and multiple channels to be handled without requiring high-end or dedicated hardware. We ensured that our solution works using commodity hardware: the various calculators involved are not required to be identical, are connected by means of ordinary LAN devices, and may be safely equipped with low-end graphics cards.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

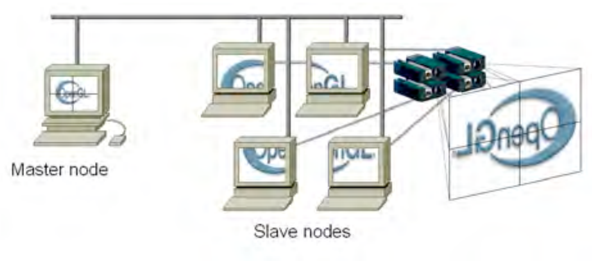


Figure 1: Cluster rendering scheme.

In our approach, each slave program is directly connected to a visualization device, so that it does not have to send the rendered image back to the master node.

The Network Renderer is completely transparent to the original XVR application running on the master node. It can be seen as a virtual OpenGL context with very high capabilities. The virtualization of the OpenGL context is performed by a software layer that intercepts all the graphics functions calls issued by the original application, called *master application*, and sends them to a set of programs, called *slaves*, running on the networked machines. Each slave, then, carries out its part of the rendering task.

The network renderer performs also the operations necessary to adapt the visual output to the morphology of the system, such as the distortion compensation and the blending of the overlapping areas of projection. Figure 3 shows how distortions, misalignments between the projection surfaces and overlapping areas are managed by the software, in order to obtain the proper visualization.

4. Hardware setup

As slave computers we employ two middle-end machines equipped with AMD Athlon 64 Dual core processor running at 2 GHz with 1 GB of RAM; the graphics card is a nVidia GeForce 7800 GTX with 512 Mb video memory. Each of them is connected to two projectors and handles a single channel on the whole.

The master machine is slightly different and employs a 2.21 GHz AMD Athlon 64 Dual and 512 Mb of video memory.

The screen is 3.4m x 2.3m, slightly more than a 160" diagonal. It is made of polarizing-preserving fabric, allowing the system to be used with polarized filters and glasses. In our setup, though, we use Infitec frequence-shifting filters and glasses, that do not require the projection surface to have any particular property.

For our research purposes we needed the flexibility to chose different filters type, but a similar setup with no such needs could employ a much cheaper projection surface.

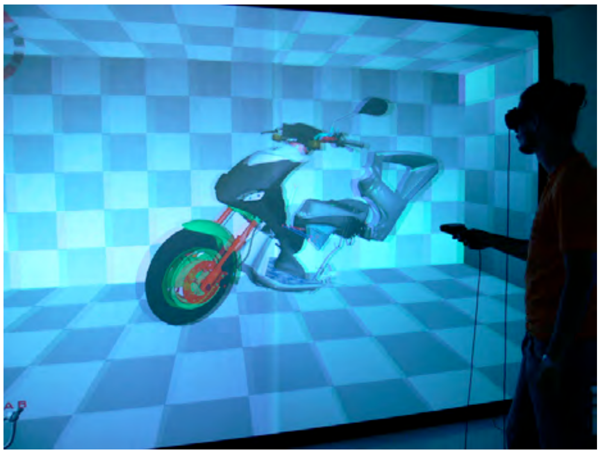


Figure 2: System developed in cooperation with Piaggio.

We employ only commodity projectors with a resolution of 1400x1050 pixels. We use pairs of them coupling them in vertical orientation in order to achieve two stereoscopic channels with a 2100x1400 resolution each. This gives approximately 15.5 dpi density.

Our tracker is a commercial Intersense IS-900 PC-Tracker. It is a 6 degree of freedom (6-DOF) motion tracking system utilizing InterSenses hybrid technology combining the output of inertial sensors with range measurements obtained from ultrasonic components.

5. Industrial design

PERCRO, in cooperation with Piaggio S.p.A., has developed a system to manage a Virtual Environment aimed to improve the assemblies quality [4]. This Interactive 3D Visualization System allows a user to visualize complex CAD assemblies in 1:1 scale, giving the impression of facing a real prototype. The user may explore the model by selecting and manipulating its components [3].

Although many users can use the system at the same time, only for one of them (the active user, whose absolute position/orientation is real-time tracked) the generated perspective is correct.

The system allows to interactively explore the assembly and to manipulate its components. It is possible to select the elements of the assembly and to move them with operations of roto-translation, to temporarily hide them in order to assist the vision of the assembly interior, and to perform measurements of distance between points of the assembly.

The choice of the interaction metaphor was one of the most complex aspects of the application. The use of a Powerwall-like system makes the traditional approach to the interaction based on keyboard and mouse not only inadequate but also very difficult to implement. For

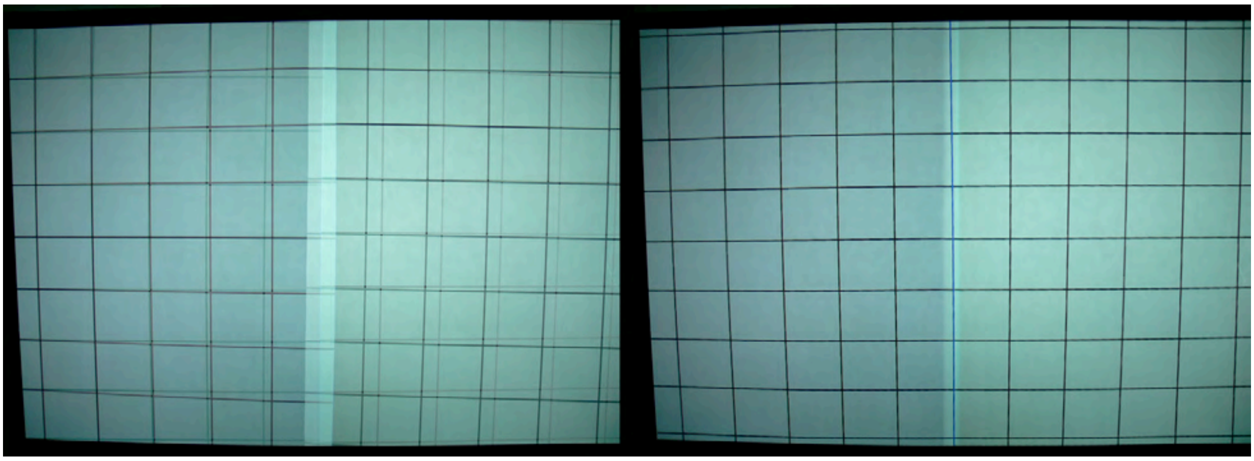


Figure 3: Output of the powerwall without (left) and using (right) software compensation

this reason it is necessary to implement an interaction metaphor which, on one side, exploits the perceptual enhancement provided by such a system and, on the other side, keeps in account the related constraints. Because the system allows the user to interact with a 3D space, the use of a 6-DOF interaction device appears the most suitable choice both for the manipulation and for the interactive exploration of the 3D model, even if an experienced CAD user may meet difficulties in getting used to these technologies so different from the traditional ones. Therefore an attempt was made to make the interaction, even if not traditional, effective and intuitive.

The basic principle was to exploit the great workspace offered by the system and the 1:1 scale factor which the virtual model is represented with. In this case it is possible to create a direct connection between the movements of the 6- DOF device (from now on: interaction controller) and the roto-translations applied to the selected components of the model (or to the whole model itself). For instance it is possible to perform a translation of a component by positioning the interaction controller in correspondence with the desired component and performing a physical translation of the device which will cause an analogous translation of the component in the virtual space.

6. Performances evaluation

The utilization of the cluster rendering architecture introduces some overhead that could affect the system performances. They could vary depending from the bandwidth that the network is able to provide. In order to improve the efficiency and to reduce the performance dependence from the network technology used, the network traffic generated by the master application is continuously compressed.

The following graphs show a preliminary analysis of the XVR Network Renderer performance. The test was performed running a master application at a 1024x768 resolution, and one slave program at a 2100x1400 resolution.

Figure 4 shows a comparison between different LAN technologies in terms of inter-frame time. It shows that achieving a resolution that is more than twice the original one introduces a noticeable overhead. Nevertheless the various curves do not differ considerably, even though they refer to very different bitrates. This is mostly because data compression results in noticeably smaller packets. As a matter of fact, if compression is disabled (Figure 5), it's easy to note that the differences between curves become greater, especially in the case of the 10Mbps link; in this case the time necessary to send packets is not irrelevant anymore if compared to network delays, and the difference increases as the bitrate decreases.

Network technology we normally employ for cluster rendering is Gigabit Ethernet.

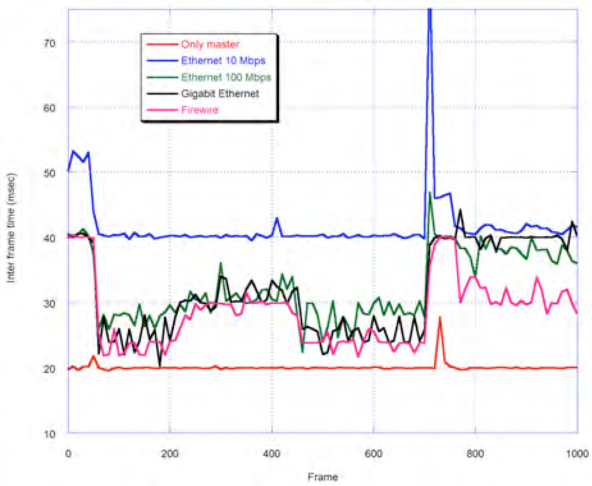


Figure 4: Test using data compression.

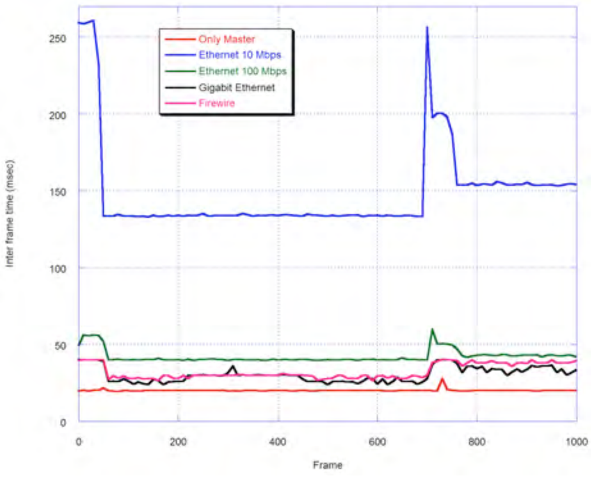


Figure 5: Test without data compression.

7. Conclusion

We have presented the software and hardware architecture of a high resolution stereo projection system based on the XVR technology, a VR software framework commonly used in the context of the ENACTIVE NoE. We have detailed the components of our system, describing in particular how the final image is composed and placed on the screen. Being our system based on a cluster of commodity PC, we have also performed and exposed some performance tests, showing what kind of numbers can be obtained depending on the cluster configuration and the underlying network technology.

The believe is that the resulting system constitute a cost-effective solution to the presentation of high-resolution and highly interactive 3D graphics content, providing a simple-to-use general architecture able to accommodate even very complex VR applications.

8. Acknowledgements

This work is partially supported by the European Commission IST-2002-002114 Enactive Network of Excellence.

References

[1] M. Carrozzino, F. Tecchia, S. Bacinelli, C. Cappelletti, and M. Bergamasco. Lowering the development time of multimodal interactive application: the real-life experience of the xvr project. In *ACE '05: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 270–273, New York, NY, USA, 2005. ACM Press.

[2] J. Jacobson. Using "caveut" to build immersive displays with the unreal tournament engine and a pc cluster. In *I3D '03: Proceedings of the 2003 symposium on Interactive 3D graphics*, pages 221–222, New York, NY, USA, 2003. ACM Press.

[3] A. Joshi and S. Rasal. Interaction paradigm for semi-literates and illiterates. In *Proceedings of 2nd International Conference on Open Collaborative Design of Sustainable Innovation*, 2002.

[4] M. Marcacci, C. Limone, M. Pieve, F. Salsedo, F. Tecchia, M. Carrozzino, A. Dettori, and M. Raspolli. La prototipazione virtuale per veicoli a 2 ruote. In *Proceedings of Convegno Nazionale XIV ADM - XXXIII AIAS, Innovazione nella Progettazione Industriale*, 2004.

[5] G. Marino, F. Tecchia, and M. Bergamasco. Cluster-based rendering of complex virtual environments. In *INTUITION 2007: Proceedings of the 4th International INTUITION Conference on Virtual Reality and Virtual Environments*, 2007.

[6] S. Molnar, M. Cox, D. Ellsworth, and H. Fuchs. A sorting classification of parallel rendering. *IEEE Comput. Graph. Appl.*, 14(4):23–32, 1994.

[7] A. Peternier, S. Cardin, F. Vexo, and D. Thalmann. Practical Design and Implementation of a CAVE Environment. In *International Conference on Computer Graphics, Theory and Applications GRAPP*, pages 129–136, 2007.

Implementation of perception and action at nanoscale

Sylvain Marlière[‡+], Jean-Loup Florens[‡], Florence Marchi[*†], Annie Luciani[‡], Joël Chevrier[*†]
(‡) Laboratoire ICA, INPG, 46 Avenue Félix Viallet, 38000 Grenoble
(+)ERGOS-Technologies, INPG, 46 Avenue Félix Viallet, 38000 Grenoble
(*)Institut Néel, CNRS/UJF, 25 Avenue des Martyrs, BP 166, 38042 Grenoble cedex 9
(†)Université Joseph Fourier - BP 53 - 38041 Grenoble Cedex 9
E-mail: joel.chevrier@grenoble.cnrs.fr, luciani@imag.fr

Abstract

Real time combination of nanosensors and nanoactuators with virtual reality environment and multisensorial interfaces enable us to efficiently act and perceive at nanoscale. Advanced manipulation of nanoobjects and new strategies for scientific education are the key motivations. We have no existing intuitive representation of the nanoworld ruled by laws foreign to our experience. A central challenge is then the construction of nanoworld simulacrum that we can start to visit and to explore. In this nanoworld simulacrum, object identifications will be based on probed entity physical and chemical intrinsic properties, on their interactions with sensors and on the final choices made in building a multisensorial interface so that these objects become coherent elements of the human sphere of action and perception. Here we describe a 1D virtual nanomanipulator, part of the Cité des Sciences EXPO NANO in Paris, that is the first realization based on this program.

1. Introduction

Nanotechnologies are already involved in cellular phones or laptops. This is nanoelectronic. This provides us with tools for massive and easy real time treatment of informations. Consequences are a new representation of space and time and revolutionized scheme of interactions among people. If ubiquity is the permanent possibility to have simultaneous interactions with different people in different places, then this is provided by nanotechnologies. Nanotechnologies will quite soon make more available. Network of nanosensors and nanoactuators will significantly extend our capacity to experience the world around us. The nanonose [1] should routinely enhance our capacity to detect different chemical species at level down to few molecules if not a single one. Medical applications such as immediate diagnostics are clearly numerous [2]. These new ways of treating information and of interacting with the matter either living or inert can be implemented in different ways using nanotechnologies. Tangible media strategies for example explore the possibility to inscribe bits of information in objects around us [3]. Our ability to physically interact with an environment is then essentially left unchanged. It is the information content of this environment that is broadly enhanced. A very different strategy, albeit not concurrent, is to extend in real time our perception and actions in the world around us [4-|6]. It is then here not reality that is augmented, but our ability to directly experience the world around us. New perception and action modalities are technologically built, based on a detailed control of how new aspects of matter organization can interact with us. Open area are numerous and are not limited to the nanoworld which however is the most challenging example. It can be amplification of minute morphological changes revealing to our senses how much grounds, houses,... are constantly vibrating, moving. One can realize how much water no longer resembles the liquid we are used to if we consider it at the micrometer scale or how much living microorganism constitution and behavior are constrained by rules that are foreign to our daily experience. One way to develop this extension of the sphere where our senses are efficient can be based on real nanosensors and nanoactuators. An another approach is to use virtual environments which can offer the nanoworld to us through real time multisensorial interfaces. This can dramatically enhance possibilities for easy exploration of remote realities foreign to our senses and can trigger a spontaneous motivation of the user similar to the one observed in video game player. Interest in these developments is for us threefold. The first two aspects are the application fields we have chosen: nanomanipulation and elementary scientific education. The third one relates to a fundamental

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

aspect of this program and roots the chosen applications. It is a major question: if the nanoworld is our chosen playground, shall we be able to define a multisensorial interface that once implemented in the nanoscene both preserves the original specificities of the nanoworld and develops the trainee ability to master this interface for an interesting and efficient use. To be relevant and successful, this program should be structured by interdisciplinary developments. The first aforementioned aspect is related to mechanical nanomanipulation of real nanoobjects. Hand on control combined with the use of sophisticated sensors and actuators is required. It is still a research project. The second aspect is more focused on new strategies for scientific education. Real time use of senses to explore "new worlds" thanks to new virtual environments can reveal aspects of matter behavior to anybody that are classically approached only through advanced scientific education.

A central part of this program is not a new idea. It goes back to birth of experimental science with use of telescope by Galileo to observe the Moon and to come to the immediate conclusion that the Moon is Earth like. As immediately emphasized by Galileo, this dramatic change in the human representation of the Universe is caused by direct use of senses technically extended by an instrument and not by a posteriori rational demonstration.

Our proposal can be seen as a revival of this famous tale. There is however a major difference. Two questions can illustrate the need for new approaches based on implementation of nanoscene in virtual environments. As nanoscale is gradually approached, continuous description no longer stands and the molecular discontinuous structure of matter is revealed. Atomic scale is a radical rupture with our common experience that is based on the objective existence of isolated continuous objects. Second is: can we offer ourselves the possibility to see and to touch an electron, a particle that has a mass and an electric charge but has no classical material spatial extension in the sense of a material sphere, although severely constrained by the Pauli principle in its collective spatial properties that are at the root of stability of matter stability. In fact seeing and touching an electron has no intrinsic meaning. Electron based objects can however be created and our interaction with these unusual objects defined. Shall we be able to build a bridge between our perception and "electron based objects" ?

If connection to the real world is considered, the simulacrum involves a numerical modelisation of the nanoworld. All measures transferred by nanosensors are translated trough real time simulation to this nanoscene, where they are experienced by the users. Symmetrically acts done by the user are both analyzed through their effects in the nanoscene where they can be simultaneously perceived for immediate feedback and transferred to the real scene for direct manipulation.

This approach is summed up in figure 1.

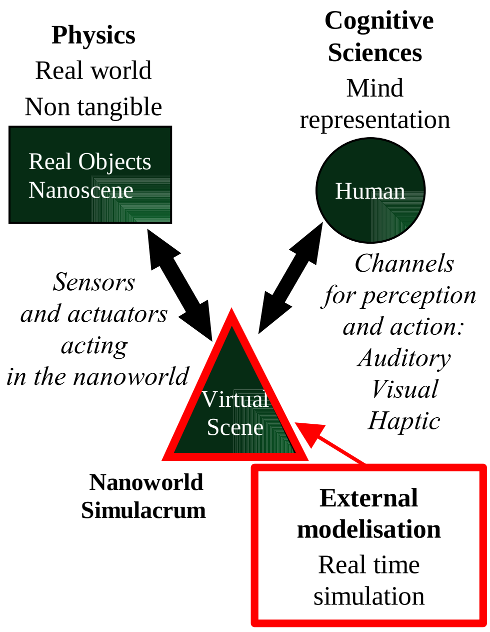


Figure 1: Basics of real time extension of perception and action to non tangible objects and foreign aspects of matter for our senses. Schematic description.

2. Feeling the nanoworld

Our point is then that seeing, hearing and touching can be technically transferred to the nanoscale. We are now testing the efficiency of this idea in scientific exhibitions. Through the use of nanosensors and nanoactuators, real time processing and visual, auditory and haptic interfaces, a multisensorial reconstruction of the nanoworld can be proposed so that one has the feeling to directly explore a foreign space. As a first step, transfer of direct repulsive contact has already been realized by many groups. Although very useful, nanomanipulators based on this, do not really challenge our perception. In this case, for example, the continuous and geometrical description of a carbon nanotube makes it appearing like a spaghetti in a plate. That we can use such a straightforward comparison is meaningful. Direct contact is finally related to the impenetrability: two solids cannot occupy the same place. They must be separated. The direct contact we are used to at our scale occurs on a distance that is finally nanometric.

Among many specific aspects, we have first chosen the universal long range interaction between nanobjects. In real life, we are not sucked by walls as we are moving close to them. At nanoscale, this is the case. This effect is often quoted as one of the most important atomic properties to understand matter structure. It immediately prevents us to define nanoobjects by intrinsic sharp geometrical contours as it is mostly always the case at our scale. Control of interaction between sensors and nanoobjects necessarily enters perception and representation. Again how can we see, touch and even manipulate with hands, for example the electron droplets created by light in CCD sensors?

3. A virtual experience : the nanocontact

The most obvious and simple way to experience interaction with a table using an instrument is to touch its surface using a stick. No need to explain the user in detail what is the situation: the user has a stick in hand and he uses it to touch the table surface coming from above. This is obvious to everybody as is then the dramatic change at nanoscale in such a simple situation introduced by the long range interaction. This is what has been proposed a general audience at public exhibition, first at the CCSTI Grenoble (Exposition Nanotechnologies) and at the Cité des Sciences in Paris (Expo Nano) and now at the Globe CERN Geneva. The stick, table surface and their interaction are numerically modeled (see figure 2) first at our scale so that, as a reference, one has the experience of the direct contact then at the nanoscale where this simple experience in fact exactly reproduces the so called force approach curves obtained using an Atomic Force Microscope.

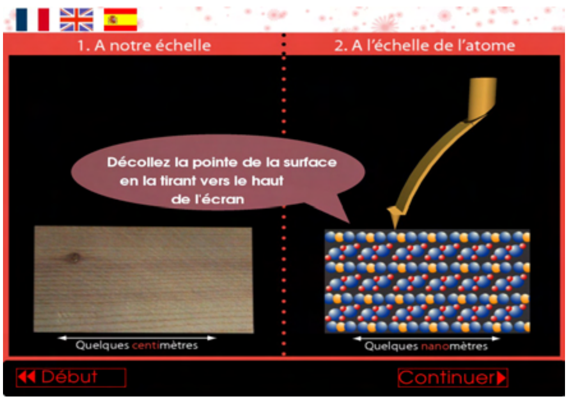


Figure 2: Copy of the screen used during the manipulation. Left: wood table at meter scale; right: atomic layer here directly touched by the manipulated elastic stick. One can continuously move from right to left and back.

One major point in this simulation is the insertion of a virtual stick that has the properties of a real one between the simulated nanoscene and the user. This stick is elastic with a major consequence as the stick end (the tip in figure 2) is getting too close to the surface, it is irreversibly sucked by the surface. It is for the user a challenge or a game very close to the actual experimental situation to stabilize the tip as close as possible to the surface without touching it. To insert a realistic tool has been done on purpose. First it carries the fact that we are not implementing perception and action in an ideal world as described by physics even though it is here a virtual world. However it must be clear that the modeled instrument characteristics, here the stick, are not due to limitations of the haptic interface. To the contrary this can be done thanks to high performance (force amplification, positioning precision, large bandwidth) of the ERGOS haptic interface.

This virtual Atomic Force Microscope is coupled to this advanced haptic interface and a sonification and visualization system (see figure 3).

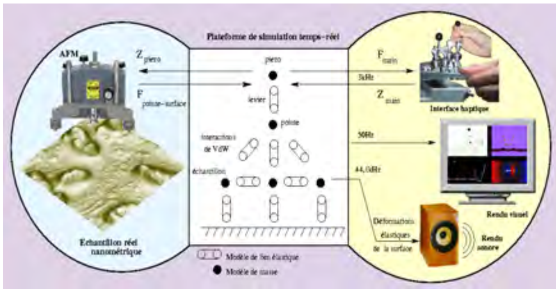


Figure 3: Image of the virtual machine. Left part: image of a real AFM; Center: elastic model of the stick, table and interaction based on mass and spring combination.

Using this machine, as expected, everybody can then compare his feelings when directly experiencing characteristics of the usual contact at our scale or specificities of the nanocontact.

We then believe we have here produced a set up that leads anybody in a few minutes to directly experience the long range interaction at nanoscale in his hands. This is finally reminiscent of the Galileo procedure: after manipulation one is supposed to be convinced about the existence and the importance of this universal long range interaction.

Figure 4: Manipulation using the virtual AFM we have built. In her hand, the haptic interface lever whose details are shown in figure 5.

One should suddenly realize the huge change it would be to invest such a world based on rules that are totally remote to our common experience in the daily environment.

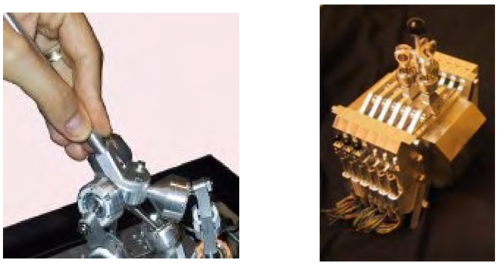


Figure 5: Details of the modular haptic interface here in a 3D configuration.

4. A preliminary evaluation

This comment is in fact our expectation. Although we have qualitatively observed this is what happening with users, this remains a weak conclusion. In order to somewhat strengthen this analysis, using inserted counters, we have analyzed the time spent by visitors on this machine at Expo Nano in La Cité des Sciences Paris.

From the March 19th 2007 to the June 1st 2007, visitor number has been 5749 i.e. about 90 visitors a day during 5 to 6 hours. The overall time on the set up spent by a user has been about 4 minutes and the manipulation time per user close to 2 minutes. From these numbers we can draw some first conclusions. This machine was offered to the public which used it freely. 90 users a day is a large number. Second an average time of 4 minutes including 2 minutes spent using simulations, appears to us reasonable to enter the system, understand the problem and get the prepared message. This result is consistent with our expextation.

This experience has put in a large audience hands a robotic set up for an extended period of time. It is then a first and encouraging step in the development of real time virtual systems equipped with multisensorial interfaces that enable one to discover aspects of our environment well scientifically described but that are totally remote to our senses. It is a research subject to observe the way people discover this surprising piece of reality and how they can get trained to it. How this training is influenced by the definition of the multisensory interface and the combination of sensory channels will be a key question. The a priori level of user scientific education needed to master the system depending on its configuration will also be interesting especially to appreciate the relevance of this procedure in scientific education.

References

[1] Sysoev V. V., Button B. K., Wepsiec K., Dmitriev S., and Kolmakov A. Nano Lett.; 2006; 6(8) pp 1584 - 1588; (Letter)

[2] http://cordis.europa.eu/nanotechnology/nanomedicine.htm

[3] Ishii, H. and Tangible Media Group, Tangible Bits: *Towards Seamless Interface between People, Bits, and Atoms,* NTT Publishing Co., Ltd., Tokyo, Japan, June 2000 (ISBN4-7571-0053-3)

[4] Marlière, S. Urma, Daniela / Florens, Jean-Loup / Marchi, Florence , *Multi-sensorial interaction with a nano-scale phenomenon : the force curve*, 2004/06/05-2004/06/07, Munchen, Germany (2004) pp.246-252 Smith, C.D., and Roberts, E.F. Book Title, Publisher, Location, Date.

[5] Marchi, F. and al., *Educational Tool for Nanophysics Using Multisensory Rendering*, 18-20 mars 2005, IEEE, Pise, Italie, vol. 1 (2005) pp.473 à 474

[6] Marlière S., Marchi F., Gay F., Chevrier J., Urma D., Florens J.-L., Cadoz C., Luciani A., 2003, *Touching nanospace : Atomic Force Microscope coupling with a force feedback manipulation system* Conference STM 03 (Scanning Tunneling Microscopy), Eidhoven, Netherlands 2003

On-line avatar control using prioritized inverse kinematics

Damien Maupu, Ronan Boulic, Daniel Thalmann
VRLab, Ecole Polytechnique Fédérale de Lausanne, Switzerland
E-mail: {damien.maupu; ronan.boulic; daniel.thalmann}@epfl.ch

Abstract

Interactive control of virtual character through full-body movements has a wide range of applications. In our approach, motion reconstruction is based on a prioritized inverse kinematics constraints solver. Several markers, placed on the user's body, act as goals for effectors. We discuss in this paper the set up we end up choosing to ensure that the reconstructed posture is not only a mathematical optimum but is also a posture that can be achieved by a real performer. Indeed, the reconstructed posture needs to correspond to the one of the user in order to have an interface as transparent as possible for the user.

1. Introduction

Interactive control of a virtual character, or avatar, has applications in fields like virtual prototyping, workspace design, and computer animated puppetry or for interacting with general purpose virtual environments. Traditional interaction devices, such as a mouse, are of limited usability in such a context because they provide too few degrees of freedom. While they are perfect for defining high level parameters to control an avatar (e.g. walking or running direction), they are inadequate to interactively define simultaneously the numerous parameters of an arbitrary posture. An enactive way to control avatars is to use one's ability to execute full body movement, i.e. through on-line motion capture (mocap).

We propose an approach to reconstruct on-line full-body (no hand or facial animation) movements using prioritized inverse kinematics. Our solution is easy to calibrate, use a minimal set of markers and is able to recover a large panel of movements from reach to locomotion.

2. Related work

Motion capture systems provide the information of position and/or orientation of sensors that can be further used to recover the posture of a human performer. In the framework of skeletal based animation, a virtual human is composed of a skeleton (i.e. a hierarchy of joints starting at a root) and a deformable skin (i.e. a mesh) that follows the underlying skeleton. Therefore, an avatar posture is defined by a vector of joint state where all joints store a local orientation and the root joint stores also the position of the skeleton root to control its global position. One key problem faced by motion capture is to determine the minimum number of sensors that allow mapping transparently the measured data into joints orientation.

Most often, the set of sensors is chosen so that combining them allows to unambiguously recover the local orientation of all skeleton rigid segments [10] (i.e. three position sensors per rigid segment). Such an approach is still in use in most motion capture studios working with optical systems, either with passive or active markers delivering a 3D position data only. The main drawback is that the markers placement takes a long time. Indeed, about forty optical markers have to be placed very precisely as explained in [7]. Badler et al. [1] have explored ways to reduce the number of sensors. They present a way to drive an avatar torso using inverse kinematics and a set of four magnetic sensors delivering position and orientation data. Molet et al. [11] proposed an analytic method to robustly distribute magnetic sensors orientation data over multiple joints. Chai et al. [6] recently introduced a motion capture technique employing video cameras and a small set of markers that interrogate a database of prerecorded movements and output the most suitable one. The main limitation is the restriction to such a database.

Prioritized Inverse Kinematics (PIK) has been used in the past to do (desktop) interactive posture edition [2]. This approach relies on a linearization of the system that may introduce some artifact as discussed in [3]. Nevertheless Peinado et al. [12] have successfully recovered a clarinet musician performance from a very partial set of six markers with the help of an additional constraint on the center of mass. In that experiment, the fact that the musician was not moving the feet was exploited to define the highest priority constraint to ensure the permanent contact with the floor. A

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

qualitative analysis of this output has revealed how to improve the balance constraint to produce a more believable motion by allowing a limited swaying of the center of mass both in the sagital and the lateral planes [9].

Besides, Boulic et al. have explored a video-based approach for on-line motion capture of the upper body posture [4]. However the nature of the input data still prevents its use for studying full body movements where the hands may get occluded. This is likely to occur in the full-body reach movement we study on in the last phase of the ENACTIVE European project [5].

As a consequence we propose in this paper a methodology to synthesize transparently and on-line a large panel of full-body motion relying on a reduced set of active optical marker using PIK. The proposed sensor set-up allows free movements including steps.

3. Prioritized Inverse Kinematics

Prioritized Inverse Kinematics allows constraints (i.e. effectors) to be associated with a priority level so that important properties are enforced first (e.g. feet stay on the ground) while less important adjustments are made in the remaining solution space [2].

The overall algorithm works as follow (Figure 1). First effectors are defined: type (i.e. constraints in position or in orientation), parent joint, set of recruited joints, priority and goal. Then in a real-time loop, motion capture data is acquired, in order to feed the effectors goals. The PIK solver finds an optimal posture variation to minimize effector errors according to their relative priority. Finally the virtual human is rendered.

Nearly each marker is used to feed a corresponding "position" effector that is located at the same place on the virtual skeleton. In some cases a few markers are grouped to control the orientation of a body part.

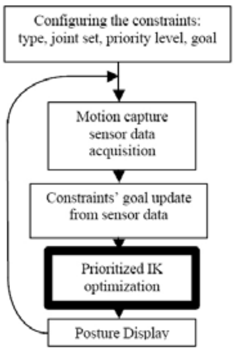


Figure 1. Exploiting Prioritized Inverse Kinematics within an on-line motion capture loop

Our mocap approach does not request markers to be carefully placed. A calibration phase similar to the one proposed by Molet et al [11] is performed once at initialization time. During that stage the performer has to adopt the H-anim calibration posture (Fig. 2, right: with the arms along the body) where all joint states are the identity [**8**]. In that state each effector and its associated marker coincide. The calibration phase allows measuring the offset distance between the effector and its parent joint in the virtual skeleton. All offsets remain constant during the on-line interaction.

4. Markers & effectors set up

Although our solution could be adapted to work with any particular mocap technology, in our current implementation we use an active optical motion capture system (8 cameras) from Phasespace [14]: markers are infra-red LEDs.

Our mocap set up (resulting from several tests) consists of 24 markers that define the goal for 19 effectors: thirteen in position control and six in orientation control.

Eight makers drive the arms. The two on the wrist help to recover the forearm orientation but does not allow recovering the wrist state (the hand is rigidly linked to the forearm as can be seen in Fig 3 top line). Only the two markers closest to the clavicles are exploited to recover shoulder shrugging movements. The spine is controlled by LEDs on the clavicles and on the spine base. Controlling the spine base orientation with a group of three markers prevents producing unrealistic curvature of the back and helps the avatar to adopt flexed knees postures. The subset of 4 lumbar and thoracic vertebras is coupled to enforce the spine anatomic behavior [15]. Legs are controlled with ten markers. The constraints on the feet have the most important priority whenever they are in contact with the floor. Otherwise, their priority is decreased on the fly during the on-line performance. This is detected when the foot markers are above a predefined height threshold. Such an approach reduces foot sliding. The head is controlled only in orientation. Joints limits inequality constraints are activated.

Figure 2 shows the user equipped with the markers. Table 1 sums up the effectors' attributes.

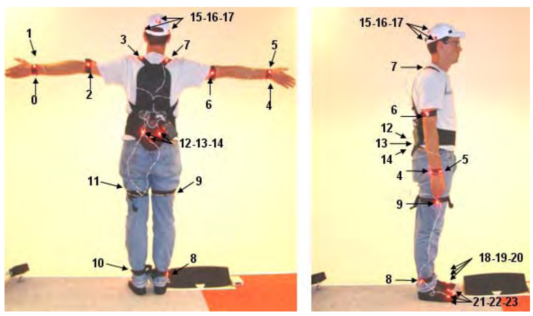


Figure 2. Setup of active markers

Table 1. Sum up of the set of constraints

Marker n°	Constraint type	Controlled body part/ recruited joint set	Priority rank
18-19-20 / 21-22-23	position & orientation	both toes / recruit joints until **Root**	**1**
12	Position	spine base / only **Root**	**2**
12-13-14	Orientation	spine base / only **Root**	**3**
0 / 4	Position	both wrists / elbow, shoulder & clavicle	**3**
0-1 / 4-5	Orientation	both wrists / only elbow twist DOF is recruited	**4**
2 / 6	Position	both shoulders / only clavicle	**5**
3 / 7	Position	both clavicles / recruit joints until **Root**	**5**
9 / 11	Position	both knees / recruit joints until **Root**	**6**
8 / 10	Position	both ankles / recruit joints until **Root**	**7**
15-16-17	Orientation	head / recruit only cervical joints	**8**

5. Results

Figure 3 presents now some results showing the large panel of possible postures our system is able to reconstruct on-line. Each iteration of inverse kinematics cost about 6 to 9 milliseconds.

6. Discussion

In the overall, the system offers a sufficiently fast convergence of the postural control so that the user feels the avatar's posture consistently reflects his or her posture over the on-line interaction.

However, there is a limitation in the frequency bandwidth of the movements. For example, if the user swings his arms fast, the avatar will not be able to track those movements accurately. This is mainly due to the local nature of the PIK solution which incrementally converges towards an optimal solution. If the position error is too large; the solver does not have sufficient iterations per displayed posture to converge towards the optimal solution.

To be able to do more iteration, we have explored a way to reduce the number of controlled variables in position control. Usually, a position control is performed in 3D to ensure correct convergence, especially when close to the target. However, when the distance from the effector to the goal is higher than a threshold, we have experimented to reduce the controlled dimension to 1D along the direction effector-goal. The computing cost being linear with the number of controlled dimensions, the gain can be significant if this happens simultaneously to multiple position effectors, e.g. during fast user movements. Although we did observe some better convergence results, it was not as significant as expected over a large range of users' activities ; it appears that the effector subset that benefit from the dimension reduction remain small in proportion of all other effectors. Also, switching the controlled dimension brings some perceptible movement discontinuities. For this reason we prefer to use this improvement only for off-line movement reconstruction rather than on-line interaction for which with favor intuitive and transparent postural control.

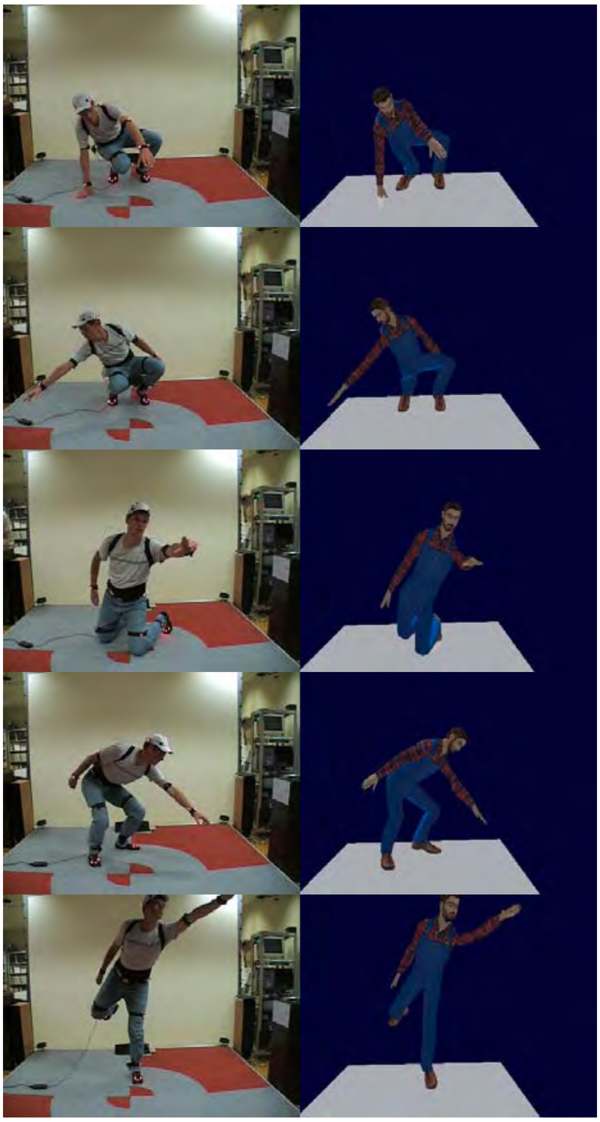


Figure 3. A large panel of reach postures can be reconstructed on-line by the system.

Another frequent improvement of the Prioritized IK convergence is to attract the current posture towards a preferred posture through the lowest priority task [2][12][3]. This is motivated by the fact that the IK solution is local and the controlled system is redundant, hence after a few minutes of interaction an unnatural postural drift may be observed [15]. Attracting the posture towards a preferred posture solves this issue if the remaining solution space is large enough. This approach makes sense whenever there is a small number of effectors (i.e. less then 5) because the solution may converge towards mathematically optimal but unnatural postures [12]. However in the present context the set of effectors is large (nineteen), hence the dimension of the remaining solution space for the posture attraction is really low. No differences were observed between an avatar attracted to a reference posture and one that is not. On the positive side, the large number of effectors ensures that no postural drift is observed either; therefore such lowest level improvement is not necessary and the related computation cost can be spared.

7. Conclusion

In this paper, we have presented a method to do on-line full-body motion capture with a minimum set of markers using a PIK solver. It relies on hard constraints (joint limits, coupled spine) and soft constraints (position and orientation effectors). Our approach is easy to calibrate and allows performing a wide range of full-body movements. The set up we introduced can be seen as an enactive interface in the sense that it allows controlling transparently a virtual human. The user doesn't need to adapt his posture to make the avatar adopt the posture he wants. The reconstructed motion reflects correctly the one of the user. However, there is a limitation in the frequency bandwidth of the movements. Such an interface will allow us to conduct other researches in the framework of virtual prototyping and explore PIK-based on-line collision avoidance as in [13].

As a future work we plan to first investigate better strategies to resolve temporary markers occlusion. In addition, Kalman filter could be a good way to smooth the orientation retrieved from groups of markers. We also plan to compare the motion we reconstruct to the one output from analytic IK solver [4] and some commercial products, such as Autodesk Motion Builder. Finally, adding a few more markers on the fingers should allow simple manual interactions such as grasping and releasing a virtual object (full-body mocap setups are usually limited for hands movements, i.e. only wrist rotation).

Acknowledgements

This research was partially supported by the european network of Excellence ENACTIVE.

References

[1] Badler, N. I., Hollick, M. J., and Granieri, J. P., "Real-Time Control of a Virtual Human Using Minimal Sensors", In Presence Teleoperators and Virtual Environments, Volume 2(1), pp. 82-86, 1993.

[2] Baerlocher, P., Boulic, R., An Inverse Kinematic Architecture Enforcing an Arbitrary Number of Strict Priority Levels, The Visual Computer, Springer Verlag, 20(6), 2004

[3] Boulic R. "Challenges in Exploiting Prioritized Inverse Kinematics for Motion Capture and Postural Control" Lecture Notes in Computer Science, Springer Berlin / Heidelberg, ISSN: 0302-9743, Volume 3881 / 2006, ISBN: 3-540-32624-3, Chapter: pp. 176 - 187 .

[4] Boulic R., Varona J., Unzueta L., Peinado M., Suescun A., Perales F. "Evaluation of on-line analytic and numeric inverse kinematics approaches driven by partial vision input", Virtual Reality 10(1), 48-61, Springer London, May 2006

[5] Boulic R., Maupu D., Thalmann D. " Considering the normalized vertical reach performance for consistently controlling virtual mannequins from full-body input", third International Conference on Enactive Knowledge, ENACTIVE06, Montpellier 20-21 Nov. 2006.

[6] Chai, J., and Hodgins, J. K., "Performance animation from low-dimensional control signals", Proc. of SIGGRAPH 2005 24, 3, 686–696, 2005

[7] CMU Mocap: http://mocap.cs.cmu.edu/markerPlacementGuide.pdf

[8] H-anim : http://h-anim.org/

[9] Mazzarino B., Peinado M., Boulic R., Wanderley M., Camurri A., "Improving human movement recovery using qualitative analysis", third International Conference on Enactive Knowledge, ENACTIVE06, Montpellier 20-21 Nov. 2006

[10] Menache A., "Understanding Motion Capture for Computer Animation and Video Games", Morgan Kaufmann Publishers Inc., 1999.

[11] Molet T., Boulic R., Thalmann, D., "Human Motion Capture Driven by Orientation Measurements", Presence, MIT Press, pp 187-203, Volume 8 (2), April 1999

[12] Peinado M, Herbelin B, Wanderley M, Le Callennec B,Boulic R, Thalmann D, Meziat D. "Towards configurable motion capture with prioritized inverse kinematics". Proc. of IVWR04, 2004.

[13] Peinado M., Meziat D., Raunhardt D., Boulic R., "Environment-Aware Postural Control of Virtual Humans for Real-time Applications", Proc. of the SAE Conference on Digital Human Modeling for Design and Engineering, July 2006

[14] Phasespace: http://www.phasespace.com/

[15] Raunhardt, D., Boulic, R., "Exploiting Coupled Joints", In Proc. of GRAPP '07, 2007

A Haptic Virtual Environment for Molecular Chemistry Education

Jun Murayama   Eric BK Lok   Katsuhito Akahane   Masaharu Isshiki   Hiromi Shimizu
Makoto Sato

Precision and Intelligence Laboratory - Tokyo Institute of Technology
E-mail: (jmy, ericbklok, m_isshiki, shimizu_hiromi, msato)@hi.pi.titech.ac.jp

Abstract

The haptic learning channel plays an important role in the learning cycle. The study of abstract concepts are better understood when hands-on activities are included. When learning about intermolecular interaction, we deal with abstract subjects and thus learning tools such as molecular visualization software and plastic molecular models are used to facilitate knowledge acquisition. These tools, however, are insufficient in assisting the understanding of force Interaction. In this research, we propose an alternative learning tool for Intermolecular Understanding that deals with those restrictions and considers the importance of the haptic learning channel. The proposed system allows the visualization and interactive manipulation of molecules. Using force feedback provided by a SPIDAR haptic device, Van der Waals and electrostatic forces are recreated. An evaluation conducted with high school students validated the positive impact of using the proposed system as a reinforcement learning tool for Intermolecular Understanding.

1. Introduction

Learning tools are developed to enhance the learning cycle. The process of acquiring knowledge is affected by a large number of variables and thus different techniques and methods are constantly developed to control some of those variables and improve the learning process.

The development of learning methods, tools, teaching dynamics, and such, should consider several learning preferences for effective knowledge transfer. Depending on the educational goal, targeted audience, available tools and/or technology, instructional designers try to create new teaching approaches for students benefit.

The constant evolvement and availability of technology has allowed the birth of new learning definitions, such as: educational technology, e-learning, computer based training (CBT), web-based learning, among others. These methods share the fact that they support the learning process through the use of various technologies. Computers, devices, media, virtual environments, networks, and such, are common technologies used, when developing interactive learning tools.

Learning tools must satisfy needs, they should be developed to enhance learning, and should be presented as an alternative that helps students acquire knowledge effectively.

Common areas where students are most likely to have difficulties understanding subjects are those where abstract concepts are often introduced. When learning physics, biology and chemistry, students are introduced to some concepts and theories, that may be difficult to understand since most of those subjects are not visible nor tangible. To solve some of these limitations, it is common to conduct activities such a laboratory practices, where students become more participative by conducting experiments and by performing activities. The reason these alternate activities are effective is due to the fact that they stimulate the haptic learning channel, which is often neglected during regular lectures (in which mostly stimulation of auditive and visual channels are executed). However, sometimes even alternate activities do not satisfy all of the limitations encountered when teaching abstract concepts. A good example, as it will be briefly explained, is found in the area of chemistry, more specifically when studying concepts in intermolecular interaction.

When learning about molecules, their structures, interactions, and such, students perform experiments, construct molecules using plastic molecular kits, balloons, sticks and so on. These activities may contribute with the understanding of molecular structures, but when dealing with Intermolecular forces and Interactions they encounter several limitations. According to Sauer et al. [8], students still have a hard time understanding the forces and interactions involved in atomic binding.

To effectively learn Intermolecular Interaction concepts and considering the importance of stimulating the haptic channel, various haptic devices have been used

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

as learning tools. In this research, we propose an alternative learning tool for Intermolecular Understanding using SPIDAR haptic device, and tested the impact of including the developed system as a reinforcement-learning tool in a regular classroom.

2. Related Work

Studies have shown that students can easily grasp abstract concepts such as scientific ideas when they are allowed to test out their hypotheses through sensorial experience-based activities. In particular, haptic sensory feedback has been used to help students gain a good understanding of scientific concepts [7][8]. For example, Reiner [6] showed that patterns of forces exerted and felt by the subjects could act as an external stimulus which helped them to construct conceptual representations close to formal physical representation of fields. Furthermore, Kilpatrick [4] In his study on the use of kinesthetic feedback as an aid to 2D and 3D force field understanding demonstrated that kinesthetic feedback improved user perception and manipulation in a simple 3D virtual world more than 3D stereo viewing did. Similarly, Brooks et al. [2] found that understanding of the binding energy of a drug molecule was much clearer through forces than via visual display in a simple 6DOF docking task.

3. System

3.1 Educational Objective

The educational goal of the proposed system is the reinforcement of concepts in intermolecular interaction studied in classrooms. Based on typical subjects taught in high schools chemistry courses, the study of intermolecular interaction focused on electrostatic and Van der Waals force Interactions.

Both interactions were recreated through the system and were felt using a SPIDAR haptic device [4]. By recreating forces using SPIDAR, we were stimulating students haptic channel which as it has been proven in the past [3] is the most preferred learning channel. The calculation of forces was based on the Chemistry At HARvard Macromolecular Mechanics (CHARMM) potential energy functions [3].

3.2 Visualization

The system's visualization followed the same rendering approach of other available molecular visualization tools. By reading data files containing information on the molecules' structure, we were capable of representing molecules in a space-filling model. The molecules were represented in space-filling model since this form of representation combined with the haptic device selected (SPIDAR) would contribute with the systems intuitive manipulation.

3D structural data obtained from Protein Data Bank files [1] was used to draw virtual molecules. The specific information taken from .pdb files consisted of the atom's identification, and the atoms xyz coordinates. Combining the atoms data with their appropriate Van der Waals radii allowed the representation of molecules on their Space-filling model using OpenGl commands. A sample of the obtained molecule's representation is shown on Figure 1.

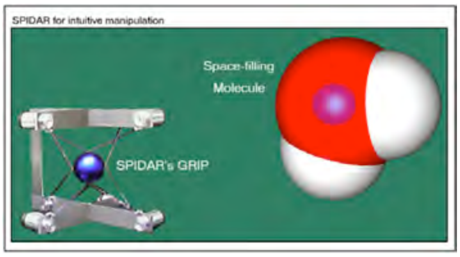


Figure 1: SPIDAR's grip- Molecule Manipulation.

3.3 Interaction

One of the goals of our research, besides contributing with students' understanding of intermolecular concepts, was to provide users with a system that could be intuitively handled. In order to have users concentrate on the educational content of the system and not on how it should be manipulated or handled, we had to provide an easy-to-use interface, and for that we selected the SPIDAR haptic device. SPIDAR provides several advantages over other commonly used haptic devices. One of those advantages is clearly shown on Figure 2, and is based on SPIDAR's grip. As shown, the SPIDAR device has a spherical grip shape. This shape

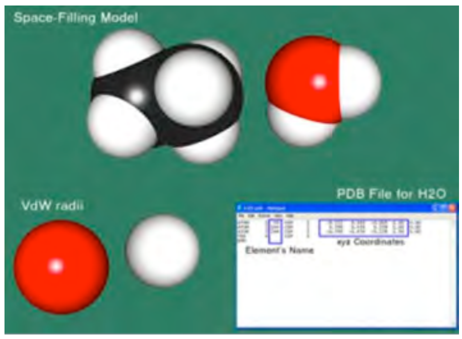


Figure 2: Molecule representation using the space-filling model.

matches the selected model in which molecules are represented. This feature contributes with the association and connection between devices grip and molecule being handled. Users can easily take control of the virtual molecule by grasping SPIDAR's grip.

System interaction consists on the movement of molecules close to or away from each other. By changing the distance between the displayed molecules and depending on the atoms involved, users can sense the realistic intermingling of repulsive and attractive fields generated from the combination of electrostatic and Van der Waals forces.

4. Evaluation

In order to evaluate the system's effectiveness as an educational tool and to test its usability, a lecture was planned and the system was tested by high school students. The system was taken to a group of students from the Salesian Polytechnic; a school located in Japan (Figure 3). The tested students were of ages 17 to 18 years old. The total sample consisted of a class of 19 students. Before scheduling the lesson with students, the system

Figure 3: Student's learning with the system.

was tested by teachers from the school. Teachers gave their impressions of the system and their suggestions on how to conduct the evaluation with the students. The lecture was conducted in the following order. Initially a lesson instructing the basis of intermolecular interactions was given to all students using traditional teaching methods, that is, a lecturer taught the class by providing students with explanations using the blackboard. Following the lessons the class was divided into two groups.

One group (Group A = 10 students) was asked to use the system right after the lesson concluded. Whereas, Group B (Group B = 9 students) took a test designed to evaluate students' understanding on the concepts taught, just right after the lecture was given. The system interaction consisted of the manipulation of water molecules. Students could sense forces between molecules by controlling SPIDAR's grip. The test consisted of six questions to inquire the students' understanding of intermolecular interactions. From these six questions, three were designed to be answered graphically. The type of molecular interactions tested is shown in Figure 4. In addition, empirical analysis via questions concerning students' impressions of the system were asked in order to evaluate the system's usability, and learner's engagement with the learning tool.

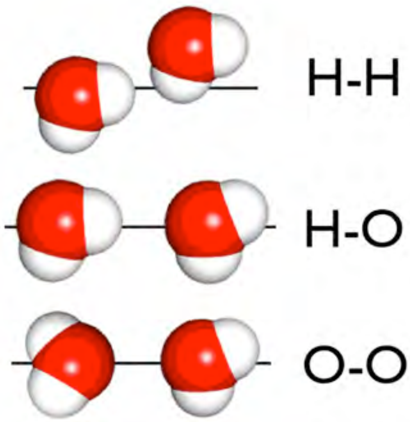


Figure 4: Molecular interactions tested.

5. Results

The evaluation conducted showed how the proposed system had a positive impact on students' understanding of intermolecular interactions. By looking at the test results from each group we notice a great difference in the number of questions correctly answered.

The difference is shown in Figure 5. It is evident how students from Group A had a better test performance. Students from Group A took the test after listening to the lecture and interacting with the system. The proposed system helped to reinforce the theory just learned and that is shown in the results obtained.

Regarding the systems general evaluation, rank questions were designed to help evaluate the system's quality. Ranks were set from one (1) to five (5). Five was considered the best option. To test the system's usability, that is whether or not the system was easy to use, the following statement was ranked according to user's experience.

Q: The system is easy to use.
(Disagree) 1 2 3 4 5 (Agree)

Users' response to that statement is shown in Figure 6. By looking at the graph, we are able to see how

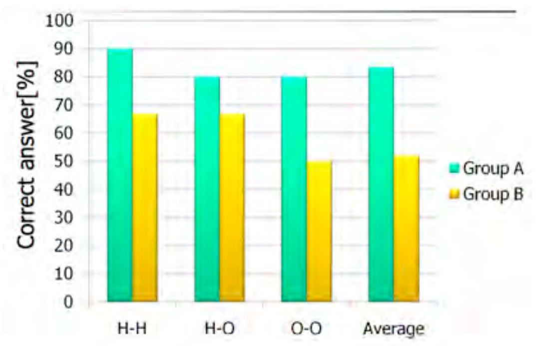


Figure 5: Results from the evaluation.

most users felt the system was easy to use and most scored the system from four (4) to five (5). Learner's engagement considers whether or not users thought using the system was fun and interesting. This was also asked using a rank question. The results show that most students enjoyed using the system (Figure 6).

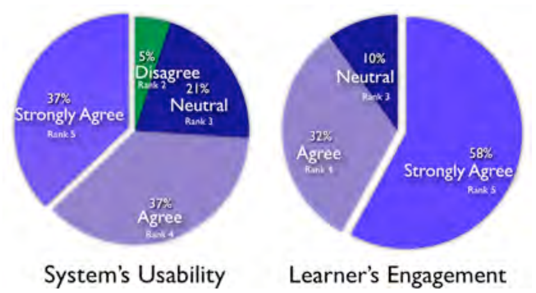


Figure 6: System's usability and learner's engagement.

In general, positive comments were received from both students and teachers. Teachers pointed out that students showed an engaging attitude towards the lesson, such that students were more interested and more enthusiastic about a lesson than regularly observed. Students made comments on how the system helped make concepts clearer and easier to understand, and how they thought the conducted activity was fun and interesting.

6. Conclusions

The system proposed for the understanding of Intermolecular Concepts using SPIDAR haptic device was proven effective as an educational tool. The system allowed the manipulation of molecules using the SPIDAR device, and intermolecular forces were sensed as molecules were moved close to or away from each other.

The evaluation of the system's functionality and system's quality conducted with students validated the important role haptic stimulation plays in the learning cycle. The results obtained showed how traditional lectures are benefited by the incorporation of interactive tools in regular classrooms.

The system targeted one specific subject (Intermolecular Interaction) for that, future development targeting other educational areas are recommended and suggested for further testing on the impact of incorporating learning tools using SPIDAR haptic device in regular classes.

References

[1] Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E."The Protein Bank." Nucleic Acids Research, 28, pp. 235-242, (2000).

[2] Brookes, F.P., Ouh-Young, M., Batter, J.J., and Kilpatrick, P.J. "Project GROPE-Haptic displays for scientific visualization." ACM Computer Graphics, 24, pp.177-185, (1990).

[3] Brooks, B.R., Broccoleri, R.E., Olafson, B.D., Stales, D.J., Swaminathan, S., and Karplus, M. "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations." J. Comp. Chem, 4, pp. 187-217, (1983).

[4] James, W.B., and Galbraith, M.W., "Perceptual learning styles: Implications and techniques for the practitioner." Lifelong Learning, 3(2), pp.20-23, (1985).

[5] Kim, S., Ishii, M., Koike, Y., and Sato, M., "Design of a Tension Based Haptic Interface: SPIDAR-G", World Multiconference on Systemics, Cybernetics, and Informatics, pp. 422-427, (2000).

[6] Reiner, M., "Conceptual construction of fields through tactile interface", Interactive Learning Environments, 7(1), pp. 31-55, (1999).

[7] Sankaranarayanan, G., Weghorst, S. Sanner, M. Gillet, A. and Olsen, A. "Role of Haptics in Teaching Structural Molecular Biology." Proceedings of the 11th Symposium of Haptic Interfaces for Virtual Environments and Teleoperator Systems (HAPTICS'03), pp. 363-366, (2003).

[8] Sauer, Ch.M., Hastings, W.A., and Okamura, A.M. "Virtual Environment for Exploring Atomic Bonding.", Proceedings of EuroHaptics 2004, pp. 232-239, (2004).

Differential Approach and Linear Motion Base Simulator Validation

Neimer Joëlle*
(*)Faculty of Psychology and Educational Sciences, University of Geneva, Switzerland
E-mail: joelle.neimer@pse.unige.ch

Mohellebi Hakim*†
(†)Renault, Technocentre CTS, TCR AVA, Guyancourt, France*
E-mail: hakim.mohellebi@renault.com

Abstract

Simulator has become an essential tool to improve our knowledge in the field of driving behavior. Unfortunately, today it remains difficult to establish the validity of such tool. The goal of this paper is to propose a new investigation method, which allows an estimate of the validity of a moving-base driving simulator. This method consists in the one hand, to identify the individual characteristics, which, in a real driving situation, lead to behavioral differences, and on the other hand, to confront the behaviors observed during simulations with those resulting from the on-road driving conditions. This method will be illustrated through a study whose first objective was to determine among several modalities of motion restitution, that which is the most similar to the real driving condition.

1. Introduction

To assess the credibility of the usage of a simulator, the question of the validity and the transferability of the results acquired on such device, is crucial. However, due to numerous limits [4] the behaviour of the virtual vehicle and performances of drivers are often very far from the ones of an actual vehicle.

Because, it is postulated that an individual will act in the same way if he is placed in similar conditions, a driving simulator will generally be considered as a valid research tool if it is established that behaviours observed within the simulator resemble those of the real life, under comparable conditions. In that case, the authors speak about a *behavioural validity* [2] which can refer to quantitative and/or to qualitative similarity, between the real and the simulated situations [5], [6].

According to us, the greatest difficulty to establish the *behavioural validity* lies on between-drivers differences. Most reports, on the individual characteristics, focus on the main effects, rather than on interactive or moderating effects of individual variables. When between-subject differences on driving are observed [16], they are not considered like potentially informative factors from the validity of the driving simulator. This is why we propose here to consider individual differences, and to see whether this potential exists and, if so, to exploit it.

1.1. Field dependence and driving performance

Individual differences, in the perception of the world, have often been related to the concept of visual field dependence-independence (DIC). This cognitive style refers to the fact that field dependent (FD) individuals are centered on visual information whereas field independent (FI) more centered on proprioceptive information. Thus, FI, more sensitive to inertial cues, would better detect inertial changes than FD.

Concerning the driving performance, it was established that FD are at a disadvantage compared to FI, in several driving tasks [1], [11], [15]. Indeed, the higher number of vehicle accidents observed from the FD, is notably attributed to the fact that they are slower to recognize the signs and to react than FI. In a real situation this difference is estimated around 0.4sec in favor of the FI.

1.2. Locus of control and driver

The Locus of Control (LOC), also considered in this study, can be defined as a personality attribute. It refers to the extent to which one feels able to influence the course of one's life (individuals with internal LOC) and to which one feels not responsible for the events that occur in one's life (individuals with external LOC) [20]. Studies [8], [14], [18] which focused on accident involvement and driving violations, have revealed that the externally oriented persons are more prone to be involved in car accidents, because they would be less careful, and would take less precaution to prevent road accidents. However, no information is given, concerning the way in which this predisposition to accidents translates itself on the level of the driving. We can suppose, in this study, that the prudent people with an internal LOC, will respect more the safety distances than external people.

1.3. Driving style

People can also be distinguished according to their aberrant driving behaviors (number of inattention, violations errors). Studies, on driving style, are mainly limited to cross cultural effects [10] and to the link with the likelihood of being involved in an accident [3], [12], [19]. The relationship between objective measures and driving style has not been directly examined. Indeed, while studies [9] demonstrated a positive relationship between sensation seeking and risky driving, others [7] showed that sensation seekers preferred shorter following distances than sensation avoiders. Like for the previous style, we can suppose that individuals with a risky driving behaviour will adopt shorter safety distances than prudent drivers.

2. Objective and Hypotheses

Our objective is to show that between-subjects differences may be considered as informative factors from the *behavioural validity* of a driving simulator. Indeed, we postulate that individual characteristics, mentioned above, will allow us to determine the way of which each participant would react in a simulated environment, if this last was in conformity with real driving situation, and in our study, to decide of the realism of various patterns of motion restitution.

From a subjective point of view: 1) If participants prefer a situation among those proposed in the experiment, one can suppose that this situation is more similar to real driving conditions than the others. 2) In addition, if FI are more sensitive to motion variations, they would made finer distinction between experimental conditions. Their subjective judgements will be more contrasted, compared to FD.

Concerning driving performance: 3) The most realistic condition should bring out differences in reaction times between FD and FI. 4) If one considers that prudent people are, all the more prudent that they do not control their vehicle, a decrease of safety behaviours could be interpreted as a better confidence towards oneself and toward the control of the vehicle, and in the case of our simulator to a better realism of the situation.

3. Method

3.1. Simulator prototype

A moving-base driving simulator (INRETS MSIS SIM^2 class), which reproduced the longitudinal, and the transitory accelerations generated on the driver chest [13], is used for the experiment. Three PC rendered and updated the visual scene projected by three Barco CRT 808S video projectors on three screens covering 156deg of the driver's visual field. A visual tilt system, which simulates the vertical motion of the visual horizon, and a sound system were also included.

For the experiment, three longitudinal displacements are chosen (without displacements, with short displacements, with long displacements), and combined with two back tilts (with or without tilt). Experimental conditions were resumed table 1.

Table 1. Experimental conditions, and their abbreviations.

		Longitudinal displacements		
		Without	Short	Long
Back tilt	Without	*(NO)*	*(N-S)*	*(N-L)*
	With tilt	*(T)*	*(T-S)*	*(T-L)*

3.2. Participants characterization

32 people participated to the experiment. All of them had a valid driver's license. Most had never used a driving simulator before and those who had used one, had only been confronted to a fixed base simulator.

Their DIC was measured by the Embedded Figures Tests [17]. Participants with extremes scores were identified as field dependent (FD) or field independent (FI), middle scores as intermediate (I). Because externality is related to risky driving [8], [18], [21] results on the LOC, considered through the Internality-Externality scale [20] and on the French version of the Driving Behavior Questionnaire [10], which measures aberrant driver behavior, are combined to distinguish prudent from intermediate and imprudent drivers.

3.3. Procedure

To eliminate participants who suffer from simulator sickness, a training phase on a highway is proposed. Such environment also allows us to avoid prior details on the experimental task, which occurs in a three lanes double way rural road (see figure 1).

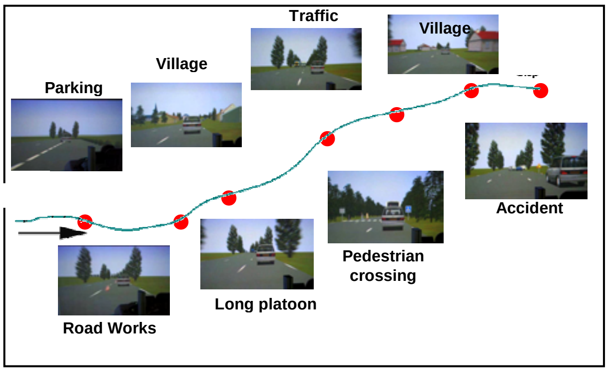


Figure 1. Example of road events

For all experimental sequences (see table 1), we implicitly forced participant to react to the lead car maneuvers. This is why this car went through a series of programmed accelerations and brake applications in a speed range from 50 to 95 km/h. To make the driving situation as ecological as possible, direct and indirect contextual cues induced the lead car velocity changes (see figure 1). An experimental condition requires that the participants drive the simulator for five minutes on average.

3.4. Recording

For each of the lead car event, five measurements (distance from the start, speed and acceleration, from the lead and the piloted cars, and intensity of the press on the braking pedal and the slow down, for the piloted car only) were recorded at the rate of 10 Hz. From all of those records Headway Time (HT), Reaction time (RT) and Time to Collision (TTC) were computed. HT is measured by taking the time that passes between two vehicles' reaching the same location; RT refers to the delay between the start of lead vehicle action and the reaction (braking or engine braking) of the driver; TTC is the time that is left until a collision occurs if both vehicles continue on the same course and at the same speed. At the end of the test, participants were required to indicate if they had preferred one condition among the six proposed, and if so, to classify those conditions.

4. Results

Two participants withdrew after the training phase because of severe simulator sickness symptoms. Their data were not used in the analyses.

From the subjective estimates, two conditions distinguish themselves from the others, the *T-S* which obtained 56% of favourable votes and the *T* condition with 40% of votes. Objective measures reveal that the RT between FD and FI differ only for the *T-S* condition. More interestingly the value of this difference is 0.44 sec in favour of the FI (see figure 2). The HT of prudent drivers are systematically higher than those of the drivers adopting a behaviour at risk, except for the *T-S* condition. Finally, it appears that prudent drivers have higher HT (1.5sec more on average), and TTC (13sec more on average) than imprudent except for *T-S*, where HT is around 2.87sec and TTC around 22sec for both groups.

5. Discussion

Our objective was to show that between-subjects differences may be informative factors, to establish the validity of a driving simulator. On the basis of on-road observations, we have considered the behaviours, according to individual characteristics and indicators of driving performance, expected in a simulated situation. We think that the reproduction, in laboratory, of between-subject differences observed in the real world, is a proof of validity of the experimental condition proposed to the drivers.

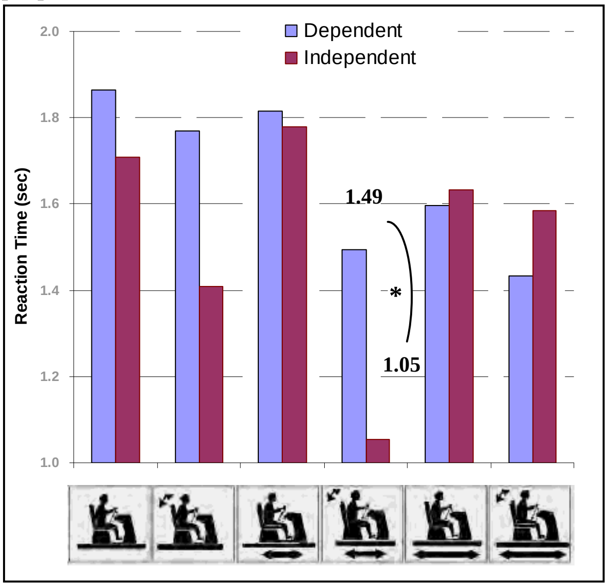


Figure 2. Interaction between experimental conditions and DIC style on the reaction times

Comparison between our expectations and behaviours observed during the simulation show that: 1) two experimental conditions seem to satisfy the participants (e.g. *T,* and *T-S*). That confirms the importance of motion feedbacks [13], without plea for a rendering of acceleration close to scale 1 [4], (long displacements did not convinced the participants). 2) Whereas we expect that FI would easily discriminate the six conditions, the finesse of their classification is not different from that of FD. This result is not alarming, and could be explained by the fact that FI have perceived the inertial differences between the conditions, but consider that they simply have the same level of realism. We would have had to add a question concerning the perception of kinetic differences. 3) We argued that a good simulator would lead to RT differences between FD and FI. Only the *T-S* condition, propose such differences, which in addition, are comparable to those observed during real driving (0.4 sec for real driving, 0.44 sec in our situation). Those two reports are clearly in favor of the validity of this situation. 4) We have also supposed that if they are confident, prudent people would take more risks during driving. It is exactly what one notes for *T-S*. Because, this condition allows participants to have a good feel of the car, we can postulate some resemblance with the real world.

What we will retain in this study, it is not the fact that we have identified a relatively valid experimental condition, which reproduces a part of the behaviors observed in real world, but rather the way in which we reached to this conclusion. Indeed, in order to question the validity of a driving simulator, we chose to exploit the existing isomorphism between the individual properties (DIC, LOC, Driving style) and the environment in which he evolves. Thus, the cutting of our population and its confrontation with various conditions of motion restitution, have allowed us to determine, what combination of motion could lead to behaviors considered as similar to those of the real world.

References

[1] Barret, G.V., Thornton, C.L., and Cabe, A. "Relation between embedded figures performance and simulator behavior", Journal of Applied Psychology, 53(3), pp. 253-254, American Psychological Association, Washington, 1968.

[2] Blaauw, G .J. "Driving experience and tasks demands in simulator and instrumented car: A validation study", Human Factors, 24(4), pp. 473-486, Human Factors and Ergonomics Society, Inc., Mesa, 1982.

[3] Dobson, A., Brown, W., Ball, J., Powers, J., and McFadden, M. "Women driver's behaviour, socio-demographic characteristics and accidents", Accident Analysis and Prevention, 31, pp. 525-535, Elservier Science Ltd , New York, 1999.

[4] Espié, S., Gauriat, P., and Duraz, M. "Driving simulator validation: The issue of transferability of results acquired on simulator", Driving Simulator Conference, North America, Proceeding, pp. 149-156, Orlando, 2005.

[5] Godley, S. T., Triggs, T. J., and Fildes, B.N. "Driving simulator validation for speed research" Accident Analysis & Prevention, 34, pp. 589-600, Elservier Science Ltd, New York, 2002.

[6] Harms, L. "Driving performance on a real road and in a driving simulator: results of a validation study". In: Gale A.G., Brown I.D., Haslegrave C.M., Taylor S. (Eds), Vision in Vehicles, Vol. V, pp. 19–26, Elsevier Science, North Holland, Amsterdam, 1994.

[7] Heino, A., Van der Molen, H. H., and Wilde, G. J. "Differences in risk experience between sensation avoiders and sensations seekers", Personality and Individual Differences, 20, pp. 71-79, Elsevier Science Ltd., New York, 1996.

[8] Hoyt, M. F. (1973). "Internal-External Locus of Control and beliefs about automobile travel", Journal of Research in Personality, 7, pp. 288-293, Elsevier Inc., New York, 1973.

[9] Jonah, B. A. "Sensation seeking and risky driving: A review and synthesis of the literature", Accident Analysis and Prevention, 29, pp. 651-665, Elservier Science Ltd, New York, 1997.

[10] Lajunen, T., Parker, D., and Summala, H. "The Manchester Driver Behaviour Questionnaire: a cross cultural study", Accident Analysis & Prevention, 36, pp. 231-238, Elservier Science Ltd, New York, 2004.

[11] Lambert, L. P., and Fleury, M. "Age, cognitive style, and traffic signs", Perceptual and Motor Skills, 78, pp. 611-624, Ammons Scientific, Missoula, 1994.

[12] Mesken, J., Lajunen, T. and Summala, H. "Interpersonal violations, speeding violations and their relation to accident involvement in Finland", Ergonomics, 45, pp. 469-483, Taylor & Francis, London, 2002.

[13] Mohellebi, H. "Amélioration des retours haptiques sur simulateur de conduite", Unpublished thesis, INRETS Paris, 2005.

[14] Montag, I., and Comrey, A. L. "Internality and externality as correlates of involvement in fatal driving accidents", Journal of Applied Psychology, 72, pp. 339-343, American Psychological Association, Washington, 1987.

[15] Ohta, H. "Individual differences in driving distance headway". In: Gale A.G., Brown I.D., Haslegrave C.M., Taylor S. (Eds), Vision in Vehicles, Vol. V, pp. 91-100, Elsevier Science, North Holland, Amsterdam, 1993.

[16] Olson, P. L. "Aspects of driving performance as a function of field dependence". Journal of Applied Psychology, 59(2), pp. 192-196, American Psychological Association, Washington, 1974.

[17] Oltman, P. K., Raskin, E., and Witkin, H. A. "Group Embedded Figures Test", Consulting Psychologists Press, Palo Alto, 1971.

[18] Phares, E. J. "Locus of Control". In: Londan, H. & Exner, J. (Eds.), Dimensions of Personality, pp. 263-304, Wiley-Interscience, New York, 1978.

[19] Reason, J., Manstead, A., Stardling, S., Baxter, J. and Campbell, K.L. "Errors and violations on the roads: A real distinction?", Ergonomics, 30, pp. 1315-1332, Taylor & Francis, London, 1990.

[20] Rotter, J. B. "Generalized expectancies for internal versus external control of reinforcement", Psychological Monographs, 80, American Psychological Association, Washington, 1966.

[21] Williams, A. F. "Factors associated with seat belt use in families", Journal of Safety Research, 4, pp. 133-138, National Safety Council and Elsevier Ltd, New York, 1972.

Enactive knowledge with gestural annotations

Jérôme Olive[*] Indira Thouvenin[*†] Stéphane Aubry[*†]
(*)Continental France SNC
(†)Laboratoire Heudiasyc, Université de Technologie Compiègne, France*
E-mail: jerome.olive@conti.de, indira.thouvenin@utc.fr, stephane.aubry@utc.fr

Abstract

The question of knowledge integration in virtual environments will be a challenge for the next few years for training, design, simulation and architecture… Many knowledge registrations such as symbolic representations (icons, cognitive maps, visualisation metaphors) are present in these spaces for abstract knowledge that we can capitalise. But the gesture is difficult to capitalise other than by its acquisition and complete description. We propose a model of gestural annotation, produced from a 3D annotation model. This model allows the user to explore a virtual environment and gives him the opportunity to constitute his own knowledge by reading and searching the annotations in a knowledge base. The process of searching annotations in a large number of annotations is an enactive process. In addition, the capitalisation of gestures concerned with the virtual mock-up gives a better comprehension of complex technical systems based on time-space navigation.

1. Introduction

When navigating in a virtual environment (VE) some metaphors are necessary [8] to guide the user in navigation, manipulation, or exploration. In the case of a virtual mock up manipulation to access specialized knowledge, it is often difficult to understand the context, the gesture for technical applications or the sequence of gestures to perform a task. As in [11], we consider as gestures any motions that can be applied as the cause of a performance, whichever be the producing system. This is a major concern in the case of industrial applications [13] such as maintenance of complex systems. Some VE are linked with knowledge bases for design but there is a large gap between the 2D and the 3D environments which is partly solved in the MATRICS© application [4]

MATRICS© proposes a 3D annotation model that is rooted in a virtual mock-up and a system of capitalisation of these annotations with the help of a knowledge base which allows the constitution of an experience by going to and from between the virtual mock-up and the knowledge base. It is an action in the sense of Varela [16] since the user perceives the annotations and must himself act to navigate in the knowledge base to find the annotations that are closely related or those that are necessary to understanding it. This enactive process is facilitated by the perceptions of multimodal annotations and by actions that are possible within the VE.

On the other hand, the knowledge of a gesture and it capitalisation, indispensable in training situations, is not possible by the model proposed by S. Aubry.

That is why a new gestural annotation model allows the acquisition of knowledge of the movement for a precise task by giving rise to the possibility of writing and reading the gesture on a virtual mock-up.

We present in this paper the MATRICS project, then the gestural annotation model and the first stage in the implementation of this model. These stages have to be integrated in a localised cognition context.

2. 3D Annotations and virtual environment

During collaboration on a virtual object, users exchange informal knowledge about this object. It might be text, images, sounds or more complex information like shapes or gestures. There is a need to capitalise this knowledge in the virtual environment [2]. Our approach is based on collaborative product design situation [15], in which annotations are a central artefact of the collaboration [7].

Therefore, we propose to use 3D annotations, as a support of this knowledge in the virtual space [1].

2.1. Definition of a 3D annotation

As the word "annotation" is part of everyday language, several definitions using different criteria can be found in literature [6], [5]. In this paper, we will consider an annotation as a mark or a document with the following properties:

- It refers to another document (the target) and support, or is the result of an activity about this document.
- It is not dissociable from its target.
- However, it is distinct from its target.

A 3D annotation needs to have the properties described above, plus these two supplementary properties:

- The target is a 3D object.
- It is contextualized in the 3D space. Namely, it is associated to a representation, a location

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

and interaction methods in the virtual environment.

In [3] we propose a theoretical model for 3D annotations; it is based on three components: annotation form, metadata and spatialisation.

We will now focus more specifically on 3D annotation with gestures. Gestures are essentially different from 3D for two reasons: first, they include the notion of time, which is not present in 3D shapes. This notion is necessary to express movement, or an order of events.

3. Gestural annotations

In the 3D annotation model previously cited, the only notion of time present is the date of creation or modification of an annotation. A problem arises then: How to capitalize the trace of activities around the virtual mock-up?

3.1. *Towards a gestural annotation model*

What we propose to do is to augment the 3D annotation model with a notion of time which will be as strong as the notion of spacialisation. This addition allows us to capitalize on the position of a 3D entity (object or user) in time. The first knowledge of this type that we want to capitalize is the technical gesture (in particular for industries in cases of maintenance). This capitalization will consist of a systematic saving (whether it be decided by the user or not) of the object manipulations and user movements. This data will be considered as separate annotations and will be linked to an ontology of gestures allowing a semantisation of the gesture. Here, it is not the gesture that we will annotate [9], but the gesture that will be contained within the annotation. In the same manner as with the previous model, we will be able to navigate in the concepts to facilitate the navigation between the annotations. For example, in the scope of maintenance, we will be able to sort all the annotations on the basis of “unscrew”.

We would also like to capitalize this process of annotation of the virtual mock-up. Indeed, we think that the path taken by the user in a virtual environment before annotating an object contributes to the definition of the annotation. “What have I seen before annotating?”, “Which objects have I manipulated?” is the knowledge that we would like to capitalize within the annotation. We consider that each of these stages can be assimilated to an elementary annotation, each contributing differently to the global annotation. During the annotation we save each of these elementary annotations (point of view, movement, selection, manipulation).

3.2. Annotation Process

During the annotation of the virtual mock-up, the user will take a path around it. These movements are decomposed into different sets of alternative movements and captured points of view. These different points of view captured by the user are as many in the context that allows us to understand why the user wanted to annotate the virtual mock-up. The figure 1 presents this process of annotation. The path of the user starts at an initial point of view (point of view of the previous annotation or point of view of entry in the system). This is followed by a series of intermediate movements and points of view which are as many of elementary parts of the cognitive path.

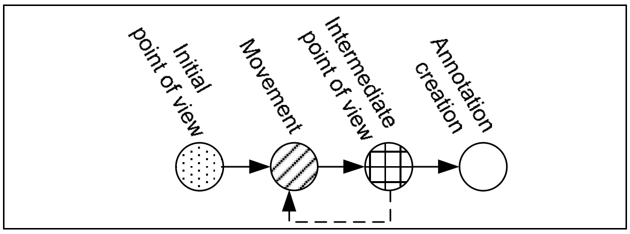


Figure 1. 3D annotation process

After having navigated (fig. 2) around the virtual mock-up, the user selects the objects to be annotated. Here two choices: either he manipulates the object in question or he annotates with classical media (text, image, sound, video…). The annotation process stops at this moment and can restart with a new annotation.

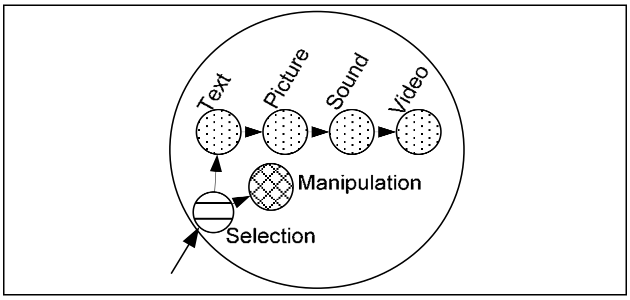


Figure 2. 3D annotation creation

Our system today allows the capitalization of annotations of movement and manipulation. We save the transformation matrices of all 3D objects (cameras, physical objects…).

3.3. Capitalization of the gestural annotation

3.3.1. *Recognition of gestures*

Thanks to a system of recognition of gestures we can give a meaning to a gesture. Our system is based on a neural network which is capable of differentiating a saved gesture from a base of pre-existing gestures. The system returns the name of the gesture which is then referenced with a base of gestures.

3.3.2. *Classification of gestures*

The identified gesture is then associated with an annotation by the creation of a link between the gesture base and the annotation. This link allows us to extract a sense of the gestural annotation and to construct a representation of this gesture in the virtual environment.

3.3.3. Gestural annotation representation

To represent annotation on the virtual mock-up, we need an anchor or an "icon": a 3D object that spatializes the annotation. This icon must show the sense of the annotation. In case of gestural annotation we need to represent the movement. In this way we studied several kinds of movement representation and notation used in choreography. Laban [10] and Benesh [12] have created two notation systems based on a human body cut-out. These two systems describe the body parts positions and movement in space. These systems have a big constraint for our needs, they describe only body movement. Sutton [14] in her early works describes the same data in her Sutton Dance Writing. But she extends her model by introducing objects of the environment and she describes the body position compared to these objects. In her latest model she describes the sign language. In this system she describes hands and face and their relative position and contact. We want to use this latest model of Sign Writing Language to create our representation for interaction between objects and human.

We propose to have a representation by annotation of the user's movement and objects' manipulation. This representation will use a system of 3D animated anchors. As in the latest Sutton system, our anchors represent interaction, action and gesture.

4. Case study

We will now study an industrial maintenance case of application. We initially consider the work of capitalization of the expert. The expert navigates around a virtual mock-up and fixes his point of view on a specific place. Then he handles an object to reach another one behind. He can now annotate the object or move it etc…. All movements are capitalized as gestural annotations. During the movement capitalization, all movements are analysed by a neural network. If a gesture is recognized only the meaning of the gesture is capitalized. On the contrary, if this one is not recognized then we preserve all the gesture data (data can be added to a gesture base with a post-treatment). In the case of a recognized gesture we adapt the shape of the annotation anchor with the meaning of the gesture.

We will now study the case of annotation reading. A user want to access to the annotation created by expert in previous paragraph. He needs to access to the carrier object before. To do that, he can read a gesture annotation (first manipulation in the previous paragraph) on the object occulting. By this action, the system replays the manipulation made by expert and access to the annotation. We think that

On the figure 3 the user needs to read annotations in red to extract screws before extract the blue piece (white annotation). When he select white annotation to extract the blue piece the annotation on screws become red to tell that they must be read.

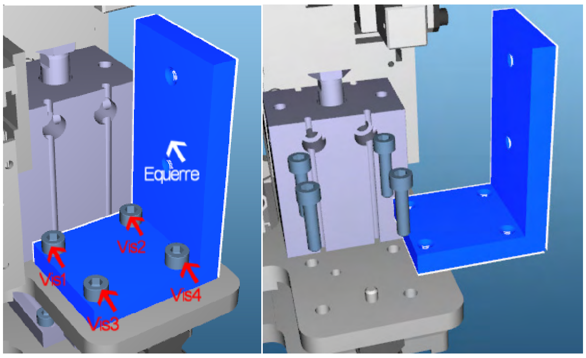


Figure 3. Read gestural annotations

5. Conclusion

In this article, we describe a new method of knowledge integration in virtual environment by an enactive process. This knowledge is registered by means of a 3D annotation model on the virtual mock-up; these annotations are capitalized by means of a knowledge base. This allows the user to construct step by step, his cognitive path. We are currently exploring a new annotation model: gestural annotation which brings localised knowledge. Our hypothesis is that this experiment when conducted in a virtual environment enriches the intellectual construction and the hard skills.

An initial system of gestural annotation was developed to capitalise the manipulations and movements made. The system will be completed by a gesture interpretation module.

In our future work, we will give autonomy to the annotations:

- automatic animation of the anchors
- presentation of the annotations' content when needed
- adaptative replay of manipulation and navigation
- evocation of the closer annotations

Acknowledgment

We want to thank Vikram Seshu for his precious and enthusiastic help.

References

[1] Aubry, S., Lenne, D., Thouvenin, I., Guénand, A., "VR Annotations for Collaborative Design", in HCII2005, Las Vegas, 22 – 27 July 2005, CD-ROM Proceedings, ISBN 0-8058-5807-5.

[2] Aubry, S., Thouvenin, I., Lenne, D., Okawa, S., "Knowledge Enhanced Annotations in Virtual reality Environments", in Virtual Concept 2005, Biarritz, 9 – 10 November 2005.

[3] Aubry, S., Thouvenin, I., Lenne, D., Okawa, S., "Knowledge integration for annotating in virtual environments", International Journal of Product Development, Vol. 5, N°6 (2007), pp. 533 - 546.

[4] Aubry, S., "Annotations et gestion des connaissances en environnement virtuel collaboratif", PhD thesis, Université de Technologie de Compiègne, 2007.

[5] Azouaou, F., Desmoulins, C., Mille, D., "Formalismes pour une mémoire de formation à base d'annotations : articuler sémantique implicite et explicite", EIAH 2003, Strasbourg, France, pp. 43 – 54.

[6] Baldonado, M., Cousins, S., Gwizdka, J., Paepcke, A., "Notable: At The Intersection of Annotations and Handheld Technologies", Proceedings of HUC Conference, Bristol, LNCS 1927, Springer Verlag, Berlin, 2000, pp. 100-113.

[7] Boujut, J-F., "User-defined annotations: artefacts for co-ordination and shared understanding in design teams", Journal of Engineering Design, vol 14, n°4, December 2003, pp. 409-419.

[8] Bowman D. A., "Interaction techniques for common tasks in immersive virtual environments", Georgia Institute of Technology, June 1999.

[9] Héloir, A., Gibet, S., Courty, N., Raynaud, M., "Système d'annotation et de segmentation de gestes de communication capturés", TALN 2005, Dourdan (France), 6-10 June 2005.

[10] Laban, R., Ullmann, L., "The Mastery of Movement", Plays, Inc. 1971.

[11] Luciani, A., Evrard, M., Couroussé, D., Castagné, N., Cadoz, C., Florens, J-L., "A basic gesture and motion format for virtual reality multisensory", in the Proceedings of the 1st international GRAPP Conference on Computer Graphics Theoryand Applications, ISBN: 972-8865-39-2, Setubal (Portugal), March 2006.

[12] McGuinness-Scott, J., "Movement Study and Benesh Movement Notation: An Introduction to Applications in Dance, Medicine, Anthropology, and Other Studies", Oxford Univ Pr (November 1983).

[13] Olive, J., Thouvenin, I., Lemasson, G., Sbaouni, M., "Tire Manufacturing supported by virtual environment", Proceedings of Laval Virtual 2006, april 2006, Laval-France. avril 2006, pp.159-163.

[14] Sutton, V., "Sutton Movement Shorthand; Writing Tool For Research", National Symposium on Sign Language Research & Teaching, Chicago, Illinois, 1977, USA

[15] Thouvenin, I., Abel, M-H., Ramond, B., Qamhiyah, A., "Environment Improvements for a Better Cooperation in Multi-Culture Collaborative Mechanical Design", In CSCWD 2002, Rio de Janeiro – Brazil, 25 – 27 September 2002.

[16] Varela, F., Thomson, E., Rosch, E., « L'inscription corporelle de l'esprit, La couleur des idées », Seuil, 1993, ISBN- 2-02-013492-

Are Moving Shapes Visually Recognized through Motor Simulations of Ocular Pursuits?

Gérard Olivier[*] Guy Labiale[*†]
(*)Experimental Psychology Laboratory, Nice-Sophia Antipolis University, France
(†) Research Team “Development, Cognition & Acquisition”, Montpellier University, France*

E-mail: olivierg@unice.fr, guy.labiale@orange.fr

Abstract

This paper deals with visual memory of moving shapes. During a visual recognition task, shapes moved on a computer screen, at a constant speed, and in a direction that was either similar, orthogonal or opposite to the direction of motion during learning. Results showed that correct response rate varies according to oculomotor factors: 1) the motor skill of ocular pursuit during learning; 2) the motor compatibility between ocular pursuits during learning and recognition. These data suggest that recognition of moving shapes involved motor simulation of ocular pursuits that subjects had previously executed during shape learning. More generally, these data question the role of eye movements in visual memory organization.

.

1. Introduction

The question of the precise role of action in knowledge organization is currently a matter of debate [4]. In this theoretical framework, the present paper deals with the links between visual recognition of moving stimuli and ocular pursuit. Let us recall that ocular pursuit is a continuous displacement of the eye that keeps a moving stimulus in the fovea, and whose motor control, like for the other body movements, relies on a covert anticipation of its execution. Previous studies have already shown that movement perception is linked to ocular pursuit. For example, when one visually tracks a moving stimulus, the perceived trajectory depends on ocular pursuit properties, and in particular, on the degree to which speed variations in ocular pursuit respect biological constraints [7]. The purpose of the present paper is to show that visual recognition of a moving shape is influenced by motor properties of the ocular pursuit that permitted it to be memorized. Let us recall that ocular pursuit is a polarized motor activity: in adults, ocular contact with the moving stimulus is more efficiently maintained along a horizontal trajectory rather than along a vertical one [6]. In the present experiment, we asked subjects to visually track shapes that were displayed on a computer screen for recognition and moved horizontally or vertically, in a direction that was either similar, orthogonal or opposite to the direction of motion during learning. Consequently, we controlled the changes in direction of ocular pursuits during learning and recognition: the interval between these two directions could be 0°, 90° or 180°. The general hypothesis is that recognition of a moving stimulus will rely on motor simulation of previous ocular pursuits executed during learning. Let us recall that motor simulation corresponds to an unconscious anticipation of executing a body movement. For example, object visual perception is sometimes accompanied by motor simulation of the manual movement that would permit the subject to reach and grasp the object [2]. Recent studies have shown that subjects mentally simulated ocular movements to recognize visual shapes [3], or ocular pursuits for predicting the future position of a moving visual stimulus after partial occlusion of its trajectory [8]. If the subjects in the present experiment recognize a moving shape through motor simulation of previous ocular pursuits executed during learning, then performance should vary with the motor compatibility between ocular pursuits during learning and recognition. More precisely, performance should vary according to ocular pursuits during learning and recognition that recruit similar or different oculomotor muscles. Let us emphasize that for 90° intervals, different oculomotor muscles control ocular pursuits during learning and recognition, whereas, for 180° intervals, the same oculomotor muscles control these ocular pursuits. In other words, for 180° intervals, there is a motor compatibility between ocular pursuits during learning and recognition, whereas for 90° intervals, there is a motor incompatibility between these ocular pursuits. Consequently, the following operational hypothesis was drawn: 90° intervals should lead to worse performance than 180° intervals.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

2. Method

2.1. Participants

Seventy-nine right-handed students took part in this experiment as volunteers (Mean age = 21.9).

2.2. Materials

Twenty-four black and white drawings of mandalas were used as stimuli. They were standardized at 4 cm in diameter. First, 12 mandalas were chosen to constitute the set of targets. Then, fifty students (who did not participate to the experiment) were required to select distractors to constitute 12 pairs of stimuli, each pair being composed of a target and the most similar distractor.

2.3. Procedure

2.3.1. Learning phase

All subjects were tested individually. Each was seated in front of a computer screen placed 60 cm before the subject. Their gaze was at the level of the center of the screen. The random set of 12 targets appeared on a uniform white background. Preceded by a white screen (1 sec.), then by a fixation cross (1 sec.), each target appeared at the center of the screen. The target moved at a constant speed straight toward one of the 4 edges of the screen (10 cm in 4 sec). and then it immediately disappeared from the screen. At the end of the movement, the eccentricity of the center of the target was 9.5°. Participants were asked to visually track the mandalas and to memorize them in order to recognize them at the end of the learning session. Six targets moved horizontally (3 rightward, 3 leftward) and 6 targets moved vertically (3 upward, 3 downward). The set of 12 targets was randomly presented 3 times. During the second and the third presentations, each target moved in the same direction as during its first appearance.

2.3.2. Recognition phase

Immediately after the learning session, the set of 24 mandalas (12 targets and 12 distractors) was randomly presented in the same conditions as the set of targets during the learning session, i.e. they moved at a constant speed from the center of the screen straight toward one of the edges of the screen. The target and its corresponding distractor moved in the same direction. Subjects had to respond with one key of the computer for a target and with the other key for a distractor. The subjects were instructed to respond as quickly and accurately as possible.

The interval between the target's direction of motion during learning and the target's direction of motion during recognition was experimentally controlled as follows (see Figure 1). Among the 3 targets that moved in a same direction during learning, one moved in the same direction again during recognition (0° interval), one in an orthogonal direction (90° interval), and one in the opposite direction (180° interval)[1]. The experimental design was as follows: P79*S2*D4*I3, with P for Participants, S for Stimulus (target or distractor), D for Direction of motion (upward, downward, rightward, or leftward) and I for Interval (0°, 90°, or 180°). The computer recorded response accuracy and response times (RT), i.e. the measurement between the instance of the appearance of the mandala on the screen and the subject pressing a key on the computer keyboard.

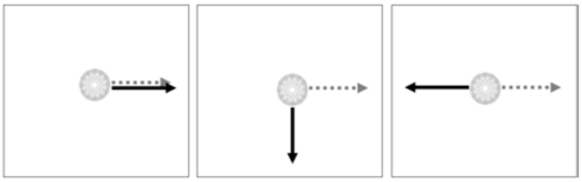


Figure 1. Example of intervals between learning (gray arrow) and recognition (black arrow) directions of motion: 0° (on the left drawing), 90° (in the middle), or 180° (on the right).

3. Results

Analyses of variance (repeated ANOVA) were performed on mean RTs and correct responses rates. RTs did not significantly vary with any factor.

For all the items, mean correct response rate was 81%. Correct response rates did not significantly vary either with Stimulus ($\underline{F}(1, 77)=1.82$, $\underline{MSE}=0.21$), or with Direction of motion ($\underline{F}(3, 231)=1.68$, $\underline{MSE}=0.10$). However, subjects gave more correct responses when the item moved horizontally ($\underline{M}$=82.6%) than when it moved vertically ($\underline{M}$=79.5%), $\underline{F}(1, 77)=4.50$, $\underline{p}$'s<.04.

For targets, mean correct response rate was 79.6%. Correct response rates varied with learning Direction of motion, $\underline{F}(3, 234)=5.18$, $\underline{p}$'s<.002. More precisely, subjects gave more correct responses when the target moved horizontally ($\underline{M}$=83.7%) than when it moved vertically ($\underline{M}$=75.5%), $\underline{F}(1, 78)=12.97$, $\underline{p}$'s<.0006. Mean comparisons showed that differences between opposite directions were not significant: correct response rates for downward ($\underline{M}$=72.9%) and upward ($\underline{M}$=78%) moving targets were not different,

[1] More precisely, after a rightward learning motion, targets moved either rightward, downward or leftward during recognition; after a downward learning motion, targets moved either downward, leftward or upward; after a leftward learning motion, targets moved either leftward, upward or rightward; after a upward learning motion, targets moved either upward, rightward or downward.

F(1, 78)=1.86, MSE=0.16, and correct response rates for rightward (M=81.4%) and leftward (M=86%) moving targets were not different, F(1, 78)=1.94, MSE=0.13.

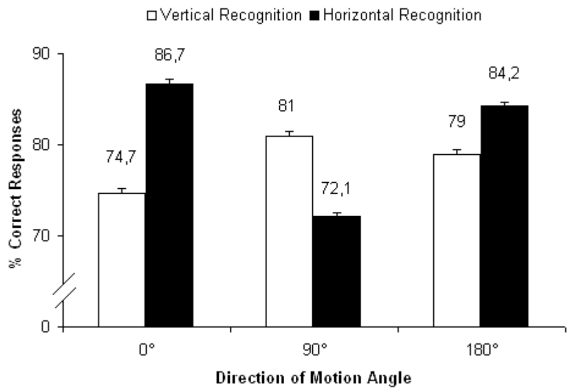


Figure 2. Correct responses rates as a function of directions of motion during the recognition task and the intervals between directions of motion during the learning and recognition tasks.

Correct response rates did not vary with Interval, F(2, 156)=1.08, MSE=0.18. However, we found a significant interaction between Interval and Direction of motion in the recognition task, F(6, 468)=2.75, p's<.02 (See Figure 2). The effect of interval was significant only when the target moved horizontally[2], F(2, 156)=5.71, p's<.005. More precisely, the 90° interval (M=72.1%) led to less correct responses than the 0° interval (M=86.7%) and the 180° interval (M=84.2%), F(1, 78)=9.98, and 7.06, respectively, p's<.01; The difference between 0° and 180° intervals was not significant, F(1, 78)=0.34, MSE=0.05. Conversely, correct responses rates did not vary with interval when the target moved vertically, F(2, 156)=1.04, MSE=0.16 (See Figure 2).

4. Discussion

4.1. Ocular pursuit polarization effect

The targets memorized during repeated horizontal pursuits were better recognized than the targets memorized during vertical ones. The greater motor skill of horizontal ocular pursuit, compared with the vertical one, explains this result. This oculomotor skill globally decreased the dual task difficulty, such that it was easier to keep the moving target stabilized in the fovea while simultaneously exploring its visuo-spatial content. Consequently, oculomotor and attentional explorations more easily structured target memorization. For similar reasons, the items that participants tracked horizontally during recognition were better recognized than the items tracked vertically. During recognition, however, the ocular pursuit polarization effect was modulated by another effect: the effect of the interval between learning and recognition directions of motion.

4.2. Direction of motion interval effect

Globally, the interval between learning and recognition directions of motion had no effect on performance. However, an interaction was observed between this interval and direction of motion during recognition: the interval influenced correct response rates when the ocular pursuit during recognition moved horizontally, but not vertically. More precisely, when the ocular pursuit during recognition moved horizontally, correct response rates for the 0° and 180° intervals were similar and above the mean, whereas the 90° interval led to a performance decrement. Let us point out that performance did not fall proportionally with the interval amplitude between the target trajectories presented during learning and recognition. Consequently, the interval effect cannot be considered as an effect of angular distance in the same manner as observed during visual mental rotation tasks. On the other hand, the interaction between interval and direction of motion during recognition can be explained by the effect of two combined oculomotor factors: the polarization of ocular pursuit during learning and the motor incompatibility between ocular pursuits for learning and recognition.

4.3. Motor incompatibility effect

The results confirmed our hypothesis: an oculomotor incompatibility effect influenced performance. According to the direction of motion during recognition ocular pursuit, this oculomotor incompatibility effect combined differently with the polarization effect of ocular pursuit during learning.

4.3.1. Horizontal recognition ocular pursuits.

During horizontal recognition ocular pursuits, the 90° interval led to a motor incompatibility effect: performance decreased because ocular pursuits during learning and recognition recruited different oculomotor muscles. On the other hand, for 0° and 180° intervals the polarization of ocular pursuits during learning improved performance (horizontal learning pursuits). Taken together, these two effects led to the global interval effect, observed regardless of the direction of motion (rightward or leftward) of the ocular pursuit along the horizontal axis.

4.3.2. Vertical recognition ocular pursuits.

[2] The same significant interval effect was found when the target moved either rightward or leftward, F(2, 156)=3.34 and 4.25, respectively, p's<.04.

During vertical recognition ocular pursuits, a horizontal ocular pursuit during learning facilitated the target memorization for the 90° interval; however, for this same interval, as ocular pursuits during learning and recognition recruited different oculomotor muscles, the oculomotor incompatibility led to a performance decrease. These two opposite effects cancelled each other. On the other hand, for 0° and 180° intervals ocular pursuit during learning were vertical and ocular pursuits during learning and recognition recruited the same oculomotor muscles: in other words, neither the polarization effect of ocular pursuit during learning nor the motor incompatibility effect influenced performance. Globally, regardless of the interval amplitude, this led to an absence of interval effect on correct response rate in vertical ocular pursuits during recognition.

4.4. Possible mechanisms

These phenomena occurred as if, during learning, the trajectory associated with each target was memorized simultaneously with the target, as already shown for memorization of rotating objects [7]. During recognition, subjects compared both visuospatial patterns (characterizing the contents of mandalas that were used as stimuli) and trajectories. A matching mechanism occurred between the perceived patterns and memorized ones in specialized cerebral pathways. However, subjects missed targets more often when the perceived trajectory during recognition did not fit the trajectory that subjects have associated with the target in memory. In other words, the target trajectory belongs to a global spatial configuration memorized by the subjects, and recognition was based on a matching mechanism between perceived and memorized dynamical configurations. The motor compatibility effect suggests that the matching mechanism between dynamical configurations actively involves oculomotor simulations. A comparison occurred between, on one hand, the ocular pursuit of the mandala actually executed during recognition and, on the other hand, motor simulations of ocular pursuits previously executed during learning of the targets. The utilization of oculomotor simulation confirms the role of motor factors during cognitive mechanisms that concern visual knowledge. The memory involved in the present experiment had an episodic dimension: subjects seem to have memorized not only a moving target, but also more globally, an episode during which they visually tracked the moving target.

To conclude, the number of moving targets that subjects recognized depended on motor compatibility between recognition ocular pursuit (executed) and ocular pursuit during learning (mentally simulated). By underlying the functional proximity between geometric and motor mental spaces [5], our results confirm the importance of taking eye movements into account to study visual memory organization.

References

[1] Liu, T., & Cooper, L. A. Explicit and implicit memory of rotating objects, Journal of Experimental Psychology: Learning Memory and Cognition, 29(4), 554-562, 2003.

[2] Olivier, G. Visuomotor priming of a manual reaching movement during a perceptual decision task, Brain Research, 1124, 81-85, 2006.

[3] Olivier, G., & Juan de Mendoza, J. L. Generation of oculomotor images during tasks requiring visual recognition of polygons, Perceptual and Motor Skills, 92, 1233-1247, 2001.

[4] O'Regan, J. K., & Noë, A. A sensorimotor account of vision and visual consciousness, Behavioral and Brain Sciences, 24(5), 939-1031, 2001.

[5] Poincaré, J. H. La valeur de la science, Flammarion, Paris, 1905/1970.

[6] Rottach, K. G., Zivotofsky, A. Z., Das, V. E., Averbuch-Heller, L., Discenna, A. O., Poonyathalang, A. & Leigh, R. J. Comparison of horizontal, vertical and diagonal smooth pursuit eye movements in normal human subjects, Vision Research, 36(14), 2189-2195, 1996.

[7] Viviani, P., & Stucchi, N. Biological movements look constant: Evidence of motor perceptual interactions, Journal of Experimental psychology: Human Perception and Performance, 18, 603-623, 1992.

[8] Wexler, M. & Klam, F. Movement prediction and movement production, Journal of Experimental Psychology: Human Perception and Performance, 27(1), 48-64, 2001.

Do the benefits of wearing spike insoles persist after taking them off in elderly?

Estelle Palluel* Vincent Nougier* Isabelle Olivier*
Université J. Fourier, Laboratoire TIMC, CNRS, Equipe « Santé, Plasticité, Motricité », UMR 5525, F-38000 Grenoble, France

E-mail: estelle.palluel@ujf-grenoble.fr

Abstract

The purpose of the study was to explore the contribution of plantar cutaneous inputs induced by a spike support surface to the control of stance in elderly. Nineteen healthy elderly (mean age = 68.8) and seventeen healthy young adults (mean age = 24.3) were instructed to stand or to walk five minutes with sandals equipped with spike insoles. Postural control was evaluated during unperturbed stance: (1) before putting the sandals (t_1), (2) five minutes after standing or walking with them (t_2), (3) immediately after removing them (t_3) and (4) after a sitting rest of five minutes (t_4). Although a significant improvement was observed in both populations after wearing them, the benefits were immediately lost after taking them off. Our results suggest that a continuous stimulation may be needed to observe long-term postural enhancement.

1. Introduction

A variety of sensory sources, including visual, proprioceptive and vestibular information, contribute to the overall control of posture and gait in humans. It is now well established that plantar cutaneous inputs also participate to balance control [6, 8, 13]. Different types of mechanoreceptors (plantar-surface (PS) and deep receptors) are involved and are widely distributed under the foot sole [7]. As the feet interface directly with the ground, cutaneous cues provide very detailed spatial and temporal information about the support surface properties and the variations of pressure under the feet that directly result from a shift of the centre of foot pressure (CoP) displacements [9]. The plantar sole is a “dynamometric map” [5].

Experimentally, cooling [3, 4], anesthetising [10, 11] or ischeming [2] the foot soles leads to a degradation of stability. Clinically, patients suffering from peripheral neuropathy (e.g., diabetics) generally exhibit an increase of postural sway. Several studies also demonstrated that tactile sensation is age- and location-related [14, 17]. By comparing young and old adults, Perry [14] concluded that both vibratory and touch detection thresholds decline with age: The loss of cutaneous sensation appeared to correlate with an impaired control of balance and an increased likelihood of falling. Other authors, by examining the effects of tactile stimulation, showed that applying vibration [15, 16] or rotary plantar massages [1] to the foot soles enhances balance control. Placing a raised edge underneath the perimeter of the plantar foot surface also facilitates postural stability [8].

Whereas reducing/suppressing or stimulating the plantar afferents are two relevant methods to explore the role of tactile messages in postural control, a third approach consists of changing the characteristics of the supporting surface. Previous findings provided evidence that standing or walking with sandals equipped with spike insoles can contribute, at least temporarily, to the improvement of unperturbed stance in elderly with relatively intact plantar cutaneous sensation [12]. The spike insole can also be considered as an enactive interface between human and his/her environment. As daily activities include standing, walking and also resting periods, the purpose of this study was to determine whether the benefits are lost immediately after taking off the sandals or whether they remain over a longer period of time.

2. Method

19 healthy elderly (8 men and 11 women; mean age = 68.8, range: 61-80 years; mean height = 165 ± 2 cm; mean weight = 73.8 ± 1.4 kg) and 17 healthy young adults (7 men and 10 women; mean age = 24.3, range 21-32 years; mean height = 172 ± 2 cm; mean weight = 66.8 ±1.3 kg) volunteered for this study. Each participant was exposed to (1) a standing and (2) a walking session. In both sessions, postural responses

were assessed during unperturbed stance with participants standing on a force platform. Three trials of 32 s with 15 s of standing rest in-between were performed (1) before putting the sandals equipped with spike insoles (t_1), (2) five minutes after standing or walking with them (t_2), (3) immediately after removing them (t_3) and (4) after a sitting rest of five minutes (t_4). The footwear consisted of the Arena® NewMarco sandals (designed for pool activities). The entire insole was covered with an array of spikes made with semi rigid PVC. As the trials were always done with the sandals, thin and flexible insoles were placed into the sandals between t_3 and t_4 to avoid the cutaneous contact with the spikes.

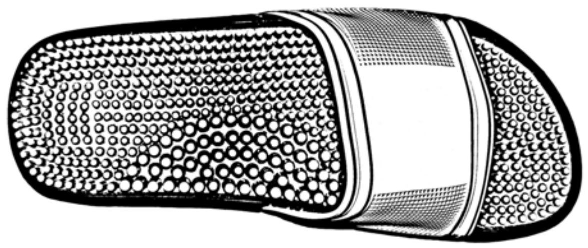


Figure 1: spikes sandals used for the study

2.1. Dependent variables

The centre of pressure (CoP) motion was processed through surface area, mean speed, root mean square on the antero-posterior (AP RMS) and medio-lateral (ML RMS) axes.

2.2. Statistical analysis

A 2 ages (young and old adults) * 2 sessions (standing and walking) * 4 times (t_1, t_2, t_3 and t_4) analysis of variance (ANOVA) with repeated measures on the last two factors was applied to determine whether the benefits of wearing spike insoles remained after taking them off. Adjustments of the p-values for the violation of the sphericity assumption were done with a multivariate test (Hotelling-Lawley Trace). Post-hoc analyses (Tukey HSD) were used whenever necessary. The level of significance was set at $\alpha=0.05$.

3. Results

A main effect of age was observed for the mean speed and the ML RMS. (p=0.001 and p=0.006, respectively), with the lower values occurring in the young adults. The analysis also indicated a main effect of time for the ML RMS (p=0.001): Whatever the age or the session, lower values of the ML RMS were obtained after wearing the sandals five minutes (p=0.045). But the benefits were lost immediately after taking them off (p=0.001) and after a rest of five minutes (p=0.001). The three-way interaction of age*session*time was significant for the CoP surface (p=0.028, see Figure 2) and the AP RMS (p=0.001).

In the elderly, as expected, post-hoc showed an improvement of the CoP surface and the AP RMS after standing five minutes with the spike insoles (p=0.001 and p=0.048, respectively). When they were removed, the benefits were immediately lost (p_s=0.001) and could not been observed after five minutes of rest for the CoP surface (p=0.001). In the *walking* session, the decrease of the CoP surface was not significant (p=0.26) but higher values appeared immediately after removing the insoles (p=0.001).

In the young adults, post hoc indicated a small improvement of postural stability in the *standing* session. There was no significant effect for the CoP surface (p_s>0.55) and the AP RMS (p_s>0.99). In the *walking* session, as expected, the decrease of the CoP surface area and the AP RMS were significant (p=0.028, p=0.030, respectively) but the benefits disappeared at t_3 and t_4 (p_s=0.001).

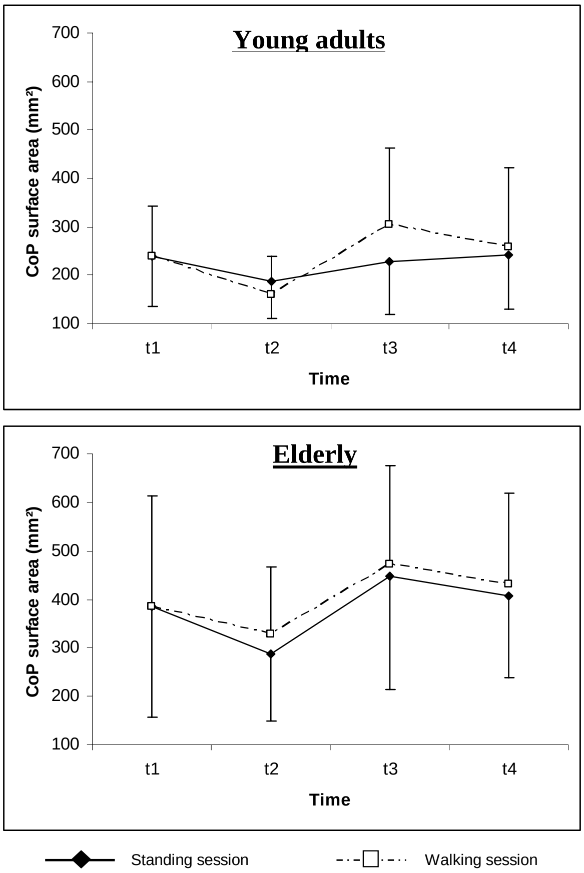


Figure 2: evolution of the surface area before putting the sandals equipped with spike insoles (t_1), five minutes after standing or walking with them (t_2), immediately after removing them (t_3) and after a sitting rest of five minutes (t_4) in young adults and elderly.

4. Discussion

The results indicated that the benefits observed on postural control in young adults and elderly after wearing the spike sandals five minutes were immediately lost when they were taken off whatever the session (standing or walking). A previous study suggested that standing or walking on a spike surface provided relevant tactile information about body position in reference to verticality: These spikes were another indented surface that increased the body awareness and improved the spatial representation of the pressure distribution under the foot soles. Therefore, they may be a relevant enactive interface in the perception field. As no correlation between PS sensitivity and postural control was found, it was assumed that the spikes might have stimulated other receptors such as the deep ones [12]. In fact, as indicated by Maurer [9], PS receptors are mainly involved in the evaluation of the support surface whereas deep receptors contribute to the continuous control of CoP displacements.

The present findings suggest that there is a need to stimulate continuously the foot soles to observe and/or to maintain postural enhancement. However, as the stimulation applied was relatively short, it remains to determine whether a longer stimulation (>5 min) would involve (1) the same benefits than those observed only after 5 min and (2) a persistence of these benefits after taking off the spike sandals. Further research is thus needed to really assess the best duration of stimulation. In other words, we can wonder whether standing or walking discontinuously but regularly with this footwear may have the same effects than wearing them continuously.

References

[1] Bernard-Demanze L, Burdet C, Berger L, Rougier P. "Recalibration of somesthesic plantar information in the control of undisturbed upright stance maintenance". Journal of Integrative Neurosciences; 3(4), pp. 433-451, 2004.

[2] Diener HC, Dichgans J, Guschlbauer B, Mau H. "The significance of proprioception on postural stabilization as assessed by ischemia". Brain Research; 296(1), pp. 103-109, 1984.

[3] Eils E, Nolte S, Tewes M, Thorwestern L, Völker K, Rosenbaum D. "Modified pressure distribution patterns in walking following reduction of plantar sensation". Journal of Biomechanics; 35(10), pp. 1307-1313, 2002.

[4] Eils E, Behrens S, Mers O, Thorwestern L, Völker K, Rosenbaum D. "Reduced plantar sensation causes a cautious walking pattern". Gait and Posture; 20(1), pp. 54-60, 2004.

[5] Kavounoudias A, Roll R, Roll J-P. "The plantar sole is a 'dynamometric map' for human balance control". Neuroreport; 9(14), pp. 3247-3252, 1998.

[6] Kavounoudias A, Roll R, Roll J-P. "Foot sole and ankle muscle inputs contribute jointly to human erect posture regulation". Journal of Physiology; 532(3), pp. 869-878, 2001.

[7] Kennedy PM, Inglis JT. "Distribution and behaviour of glabrous cutaneous receptors in the human foot sole". Journal of Physiology; 538(3), pp. 995-1002, 2002.

[8] Maki BE, Perry SD, Norrie RG, McIlroy WE. "Effects of facilitation of sensation from plantar foot-surface boundaries on postural stabilization in young and older adults". Journal of Gerontology series A Biological Sciences and Medical Sciences; 54(6), pp. M281-287, 1999.

[9] Maurer C, Mergner T, Bolha B, Hlavacka F. "Human balance control during cutaneous stimulation of the plantar soles". Neuroscience Letters; 302(1), pp. 45-48, 2001.

[10] Meyer PF, Oddsson LI, De Luca CJ. "The role of plantar cutaneous sensation in unperturbed stance". Experimental Brain Research; 156(4), pp. 505-512, 2004

.

[11] Meyer PF, Oddsson LI, De Luca CJ. "Reduced plantar sensitivity alters postural responses to lateral perturbations of balance". Experimental Brain Research; 157(4), pp. 526-536, 2004.

[12] Palluel E, Nougier V, Olivier I. "Do spike insoles enhance postural stability and plantar-surface cutaneous sensitivity in elderly?", submitted.

[13] Perry SD, McIlroy WE, Maki BE. "The role of plantar cutaneous mechanoreceptors in the control of compensatory stepping reactions evoked by unpredictable, multi-directional perturbation". Brain Research; 877(2), pp. 401-406, 2000.

[14] Perry SD. "Evaluation of age-related plantar-surface insensitivity and onset age of advanced insensitivity in older adults using vibratory and touch sensation tests". Neuroscience Letters; 392(1-2), pp. 62-67, 2006.

[15] Priplata AA, Niemi JB, Harry JD, Lipsitz LA, Collins JJ. "Vibrating insoles and balance control in elderly people". Lancet; 362(9390), pp. 1123-1124, 2003.

[16] Priplata AA, Patritti BL, Niemi JB, et al. "Noise-enhanced balance control in patients with diabetes and patients with stroke". Annals of Neurology; 59(4), pp. 4-12, 2006.

[17] Wells C, Ward LM, Chua R, Inglis JT. "Regional variation and changes with ageing in vibrotactile sensitivity in the human footsole". Journal of Gerontology series A Biological Sciences and Medical Sciences; 58(8), pp. 680-686, 2003.

Bridging the gap: from cognitive sciences and traditional media to new media

Elena Pasquinelli [*]
(*)Institut Jean Nicod (EHESS - DEC, ENS – CNRS), France
E-mail: Elena.Pasquinelli@ehess.fr

Abstract

The present contribution aims at favouring a transfer of knowledge from the domains of traditional media theory (in particular cinema cognitive theory) and from the psychology of perception to new media theory.

The notion of illusion of reality is here taken as exemplary of the possibilities offered by this transfer in terms of conceptual analysis, individuation of cognitive bases of users' behaviors and ethical considerations.

1. Introduction

The notion of illusion of reality dominates the field of new media (in particular virtual and mixed reality) through the notions of presence and illusion of non-mediation. The "official concept of Presence" [10] essentially makes reference to the possibility of *deceiving* the users as for what concerns the origin of their experience: the users fail to recognize the role of technology. Multisensoriality and interactivity are widely accepted as arguments which account for the qualitative shift from the realm of traditional representation to the realm of a full illusion of reality [6], [7], [15], thus making of enactive interfaces a privileged field for discussing the conceptual, cognitive and ethical aspects of the illusion of reality as illusion of non-mediation.

First, I will attack myself to the illusion of non-mediation and I will present evidence that traditional media audiences are aware of the fictional, mediated nature of their experience.

Second, I will illustrate perceptual studies suggesting that this awareness is not contingent to traditional media, but affects mediated experiences in general in virtue of perceptual information about the medium.

Third, I will present ethical considerations that are operational in traditional media codes of ethics and that suggest that the obliteration of the difference between mediated and non-mediated, fictional and non-fictional experiences should not be taken as a goal for the new generation of virtual reality and enactive media.

Fourth, I will discuss some possibilities of naturalization of the aspect of the illusion of reality, which is termed 'illusion of transportation' or 'illusion of presence'.

2. The paradox of fiction: cognitive awareness of fiction

Philosophers of art and fiction have long discussed about the possibility of audiences and readers of being massively fooled by works of fiction of different kinds. An exemplary case of this discussion is represented by the debate about the so-called paradox of fiction [12], in which deception is equated to the fact of holding *existence beliefs* (beliefs about the real existence) about the characters represented on screen or described in books. The paradox can be so expressed: how can we be moved by the fate of Anna Karenina without believing in the existence of Anna Karenina? The debate shows a substantial accord on the fact that audiences and readers do *not* hold existence beliefs about fictional contents, and this accord is based on the observation that responses to fictional contents are similar but never identical to responses to corresponding non-fictional contents [3]. Two strategies are then adopted in order to explain how we can be moved by the fate of Anna Karenina. So-called Thought theories propose that it is sufficient that certain events are mentally represented, or entertained in thought, in order to raise emotional reactions, eventually by evoking thoughts about real people and events. Existence beliefs are hence not required for experiencing full emotional states. Pretend theories advance that emotions and other reactions that are raised by fiction are not the same than reactions evoked by real world entities and events. They are just quasi-emotions [20] or full emotions with inhibited behavioural responses [4]. In any case, these quasi-emotions and other responses are the effect of the reader's or spectator's activity of imagination. Imagination is a cognitive activity which shares many similarities with believing, but which is essentially different because it runs off-line as other forms of

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

simulation. However, the same relations that hold among beliefs hold among imaginings [4]. Through this particular relationship that holds between believing and imagining it is possible to explain both the similarity and the difference between reactions to fictional and to non-fictional experiences. For this reason, the idea that audiences and readers accomplish off-line simulations has a greater explanatory power than the idea that they hold existence beliefs (equivalent to the idea of illusion of reality).

Can this consideration be extended so as to include new media such as virtual reality and enactive interfaces? Considerations about the existence of different reactions towards experiences with simulators (including sickness) and experiences with the real world have been expressed for new media as well [17], [19]. As in the case of traditional fictional media, these divergences as well as similarities must then be explained. A model can be proposed according to which new media users' responses (motor, perceptual and cognitive) are appropriate both to the contents and to the context of the experience (double appropriateness hypothesis). Thus, users will appropriately try to move and navigate in a virtual environment, but they will comply to the constraints offered by the device (a mouse, a joystick or a treadmill). Exactly in the same way in which horror movies' spectators will feel fear but will not try to catch the killer's hand.

This model presupposes that new and traditional media users are *aware* (at the cognitive level) of the fictional nature of the experience and *informed* (at the perceptual level) of its mediated character. It presupposes as well that users perceive the characteristics of the medium and act in accord with them.

3. The perspectival nature of perception: perceptual awareness of mediation

In the debate concerning the paradox of fiction focus is on the audiences' cognitive awareness of the fictional nature of the experience (no existence beliefs, but activation of imaginative activity with internal simulation), and little consideration is devoted to *perceptual information* about mediation. However, in addition to the fact of activating specific cognitive attitudes, perceptual media (such as cinema, virtual reality devices and enactive interfaces) present specific perceptual features as well. As simulated things, material devices for simulation are in fact real things that fall under the activity of the perceptual senses. The global perceptual stimulation (termed ambient array by [18]) obtained through a medium is hence different from the global perceptual stimulus obtained through direct contact with reality, and includes the perceptual characteristics of the medium itself. In other words, both the contents and the context (the medium and its characteristics) of the experience contribute to specifying the global perceptual stimulation.

Additional perceptual information about the presence and nature of a medium can be provided by the perspectival aspects of the perceptual content. Perspectival contents of perception are those properties of the perceptual content that depend on the position of the perceiver and on her actions [11]. An insect seen through magnifying lenses or represented in a cinematic or static image, can eventually leave the factual content unmodified (the insect), but it will necessarily change the perspectival content of non-mediated seeing, in a way that is typical of the fact of using magnifying, static or dynamic images, with related possibilities of perception and effects of action. Photographs for instance, can represent reality with high fidelity, but cannot provide the same egocentric information provided by ordinary seeing (information about the relation of the depicted objects towards the body of the perceiver) [4]. Contents perceived through movies do not vary with the actions of spectators. And enactive interfaces and virtual reality devices are not capable, at least for the moment, of providing perceptual feedback in real time to all the possible actions users can accomplish [5].

4. Ethical issues raised by the illusion of reality as illusion of non-mediation

Illusions of reality can have complex ethical consequences. A world-famous attempt to produce an illusion of reality by the mean of a traditional medium is constituted by the broadcast "War of the Worlds" (from E. G. Wells), ideated and played by Orson Welles and diffused in the framework of the series Mercury Theatre on the Air in 1938 by CBS. At some extent the attempt worked (part of the audience was fooled): a certain number of American people resulted convinced of being under Martian attack, and reacted with panic. CBS escaped punishment because the fictional nature of the performance was reminded through all the broadcast (even if with large holes, for instance, between minute 12 and minute 40 of the broadcast). However, CBS was formerly invited to avoid the "we interrupt this program" device in fictional contexts. Since then, US TV broadcasts featuring realistic news bulletins post messages to inform audiences of the fictional nature of the spectacle.

The risks and side-effects of illusion of reality are hence well known in the domain of traditional documentary media, where ethical guidelines are put in place in order to avoid gullibility and deception, not only fraud [14]. It is consistent with this concern that narrative media as well tend to frame in order to distinguish it from non-fiction [1]. It is a duty for those who propose a different attitude towards new media such as virtual reality or enactive interfaces to demonstrate that the case of new media is different from the case of traditional media and how.

5. Psychological explanations for the illusion of reality as illusion of transportation

The notion of illusion of non-mediation does not exhaust the idea of illusion of reality as it is developed in the context of new media. Another aspect of the notion of illusion of reality consists in the possibility of producing a sense of transportation in the virtual world or sense of presence. This rather obscure notion can be naturalized through the reference to a number of proprioceptive illusions that are documented in psychological literature, such as: proprioceptive illusions induced by mirrors [9], proprioceptive illusions induced by prisms [8] and proprioceptive illusions induced by artificial limbs [2]. In all these cases, thank to the appropriate manipulation of visual stimuli, the subjects feel that some of their body parts are displaced in comparison with their real position. Movement or vection illusions can be induced as well, and even through the proposition of appropriate acoustic stimuli [13]. It is a common characteristics of each of these crossmodal illusions or crossmodal biases that the presence of multiple sensory modalities (touch for the artificial limb illusions and audition plus vision or vibration for vection) and the possibility for the perceiver of accomplishing some movements (especially in the case of mirror and prism illusions) enhances the chances of experiencing the illusion and its strength.

This fact would justify in a certain measure the assertion that the multimodal and interactive character of new media produces a qualitative shift in relation to traditional, non-interactive, more limitedly multimodal media.

However, evidence produced by experiments with reaching tasks shows that proprioceptive illusions have a negative effect on reaching performances [9]. This fact confirms more general considerations about the disruptive effect violations of coherence and intersensory discrepancies [16]. The aim of producing illusions of transportation should hence be attentively evaluated against the general objectives of the application. Illusion of reality might not be, even when it can be realized, a desirable effect.

6. Conclusions

The general indication that can be extracted from the debate of the paradox of fiction, the considerations about the nature of perception, and the ethical rules adopted for traditional media is the following: the production of an illusion of reality as illusion of non-mediation cannot be an objective for new media, both for epistemic and for ethical reasons. On the contrary, in addition to the sub-personal perception of the difference between mediated and non-mediated experiences, the awareness of the fictional nature of mediated experiences should be favoured. However, aspects included in the notion of illusion of reality, such as the illusion of transportation or illusion of presence can find empirical justifications in the existence of local, specific proprioceptive illusions.

References

[1] Bordwell, D. (1985). *Narration in the Fiction Film*. Madison: University of Wisconsin Press. 18

[2] Botvinick, M., & Cohen, J. (1998). Rubber hands "feel" touch that eyes see. *Nature Reviews Neuroscience, 391*, 756. 21

[3] Carroll, N. (1990). *The Philosophy of Horror; or, Paradoxes of the Heart*. New York: Routledge 6 7

[4] Currie, G. (1995). *Image and Mind: Film, Philosophy, and Cognitive Science*. Cambridge: Cambridge University Press. 9 10 15

[5] Dennett, D. C. (1991). *Consciousness explained*. Boston: Little, Brown & Company. 16

[6] Dinh, H. Q., Walker, N., Song, C., Kobayashi, A., & L.F., H. (1999). *Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments.* Paper presented at the IEEE Virtual Reality.

[7] Durlach, P. J., Fowlkes, J., & Metevier, C. J. (2005). Effect of variations in sensory feedback on performance in a virtual reaching task. *Presence: Teleoperators and Virtual Environments, 14*(4), 450 - 462.

[8] Hay, J. C., Pick, H. L., Jr., & Ikeda, K. (1963). Visual capture produced by prism spectacles. *Psychonomic Science, 2*(215-216). 20

[9] Holmes, N. P., Crozier, G., & Spence, C. (2004). When mirrors lie: "Visual capture" of arm position impairs reaching performance. *Cognitive, affective & behavioral neuroscience, 4*(2), 193-200. 19 23

[10] International Society for Presence Research. (2000). *The Concept of Presence: Explication Statement.* Retrieved <03/2007> from http://ispr.info/.

[11] Noe, A. (2003). Causation and perception. *Analysis, 63.2*, 93-100. 14

[12] Radford, C. (1975). How Can We Be Moved by the Fate of Anna Karenina? *Proceedings of the Aristotelian Society, 49*, 67-80. 5

[13] Riecke, B. E., Schulte-Pelkum, J., Caniard, F., & Bulthoff, H. H. (2005). *Spatialized auditory cues enhance the visually- induced self-motion illusion (circular vection) in Virtual Reality*. Tubingen: Max Plank Institute. 22

[14] RTNDA, (2000). Code of ethics and professional conduct. *http://www.rtnda.org/ethics/coe.shtml*. 17

[15] Schubert, T., Friedmann, F., & Regenbrecht, H. (1999). *Decomposing the sense of presence: Factor analytic insights.* Paper presented at the 2nd International Workshop on Presence, University of Essex, Colchester, UK. 2

[16] Stein, B. E., & Meredith, M. E. (1993). *The merging of the senses*. Cambridge, Mass.: MIT Press. 24

[17] Stoffregen, T. A. (1997). Filming the world. An essay review of Anderson's The Reality of Illusion. *Ecological Psychology, 9*, 12, 161-167. 12
[18] Stoffregen, T. A., & Bardy, B. G. (2001). On Specification and the Senses. *Behavioral and Brain Sciences, 24*(1). 13
[19] Stoffregen, T. A., Bardy, B. G., Smart, L. J., & Pagulayan, R. J. (2003). On the nature and evaluation of fidelity in virtual environments. In L. J. Hettinger & M. W. Haas (Eds.), *Virtual & Adaptive Environments: Applications, Implications, and Human Performance Issues* (pp. 111-128). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. 11
[20] Walton, K. (1978). Fearing Fictions. *Journal of Philosophy, 75*(1), 5-27.

Enaction and the Ethics of Simulation

Simon Penny
University of California Irvine
E-mail: penny@uci.edu

Abstract

The goal of this paper is to enhance critical discussion of interactive media practice and interactive media cultural practice by introducing a consideration of the implications of enactive embodied involvement in the process. Out of this arises a question of ethical responsibility regarding the way cultural objects might function as training environments to build behaviors which will ultimately be expressed in the real world.

1. Introduction

While sociologists and anthropologists have examined virtual communities, gaming culture and related cyberspatial phenomena, interest has centered upon issues of identity, subjectivity and community. The evaluation of the psychological and sociological aspects of interactive entertainment has, to my mind, been limited. The embodied, enacted dimension of interactive entertainment has not been adequately considered. In particular, embodied interaction with a representation, where bodily action changes the representation in a way which is analogous to, and is designed to be analogous to, human action in the world of physical objects and forces, raises scenarios which conventional critiques of representation, and those aspects of art theory which remain influenced by traditional psychology of visual perception, are not well equipped to deal with.

The core of this conversation then is in the space between pictorial representation and simulation, or rather, in the grey and murky area where they overlap. We need a new way to think about the relation between user behavior and digital representations in interactive entertainment. The embodied aspects of simulation feed back onto representation, make representation not inert but interactive. In order to gain some purchase on this territory I want to juxtapose three aspects of human activity: interactive entertainment, professional simulator training and the not-technologically-facilitated learning of bodily disciplines and regimes of behavior.

2. Body training

One need go no further that Foucault for persuasive argument and evidence that bodily training is a powerful tool in the formation of citizens [1]. Repetitive physical actions have been an integral part of education and socialisation since pre-literate times. Anthropological observations by Marcel Mauss (and others) attest to the unacknowledged but pervasive power of physical behaviors in social and cultural formation [2]. Indeed, physical imitation is a key component in social development. The establishment of gender roles, for instance, through such emulation, voluntary or coerced, is well documented. Pierre Bordieu and others have established that social behaviors are often learned without conscious intellectual understanding [3]. The way someone rationalises or explains an activity on an intellectual level, and the behaviors which have been learned and are enacted can be different, even diametrically opposed. There is a small industry of corporate training in the reading and deployment of body language. Legion are the training and rehabilitation systems which rely on repetitive physical action, even to exhaustion. From this perspective military boot camp, football training, some forms of yoga and other spiritual training, ballet lessons and some schools of drug, and juvenile delinquent other psycho-social rehabilitation are almost indistinguishable, except in the level of academicisation of the particular techniques. One quality which is common to sports training, martial arts training and military training is anti-intellectuality. Whether an activity is introduced verbally and methodically or is instilled by discipline and repetition, it is universally acknowledged by both teacher and (successful) student that the training it is only really effective when it becomes automatic, reflex. It becomes not conscious. The unsuccessful student is told 'you think too much'.

3. The Military Entertainment Complex

Computer simulated immersive environments are clearly an effective tool for bodily training, demonstrated by their use in civil aviation and in the military. Over the last decade, applications have broadened, VR simulations have even been applied to psychotherapy. Such simulations create a useful environment for desensitising phobic patients who transfer what they've learned in the 'simulated' world to the 'real' world, allowing them to ride elevators and cross high bridges. So while the electronic game industry vehemently counters claims that interactive electronic games have any real-life consequences, psychotherapists employ simulation technologies precisely because they have effect in people's lives.

Training simulation and interactive entertainment were born joined at the hip. There is no better place to examine that join than SIGGRAPH, the Special Interest Group in Computer Graphics and Interactive Techniques of the Association for Computing Machinery (ACM). Here the military simulator development community, the academic computer science community and the high end civilian computer graphics and animation community blend. For in truth, there is substantial overlap, and personnel movements between the communities is constant and smooth.

During the 1980's, DARPA (the Defense Advanced Research Projects Agency) and the US Army developed simulators for tanks and vehicles which were integrated in a local network, a multi-user virtual environment. About the same time that Simnet became public in the early 90s, a high end immersive multi-user battle simulation game called Battletech opened in Chicago. In late1999, the the University of Southern California received a five-year contract from the U.S. Army to establish the Institute For Creative Technologies (ICT). The ICT's mandate is: "to enlist the resources and talents of the entertainment and game development industries and work collaboratively with computer scientists to advance the state of immersive training simulation". [4] The fact that the US military have invested millions in Simnet, STOW and other simulation training systems is proof enough that simulation is an effective tool for such training. It is clear that immersive simulation environments are effective in producing such training, and that such training transfers usefully to the 'real world'.

In the mid 90's, it was revealed that the US Marines had licensed Doom from Id Software and built "Marine Doom," to use as a tactical training tool. We are drawn to the conclusion that what separates the first person shooter from the high end battle simulator is the location of one in an adolescent bedroom and the other in a military base. And having accepted that simulators are effective environments for training, we must accept that so too are the desktop shooter games. The question is: what exactly is the user being trained to do?

David Grossman, a retired Lieutenant-Colonel and expert at desensitizing soldiers to increase their killing efficiency is well known for his opposition to violent video-games on the basis that the entertainment industry conditions the young in exactly the same way the military does, they hard-wire young people for shooting at humans.[5] On the other hand, advocates for game culture do their best to downplay such associations. While game environments have advanced markedly in graphical, narrative and social complexity in recent years, the 'first person shooter' remains a popular class of game, with little enactive development from the paradigmatic Quake and Doom.

Is it unfair to blame such atrocities as the Columbine and Jonesboro school massacres purely on such products? Clearly most people, even most Quake players, have a reasonable grasp of the difference between simulation and real life. But equally clearly, these games would not find a market if a larger cultural formation had not prepared the ground. It is in this context that we must ask: what behaviors do these games train? Grossman is explicit, not just about their desensitising role but about their ability to efficiently build killing skills: "Whatever you train to do, under stress, is coming out the other end. That's why we do fire drills. That's why we do flight simulators...Well, when the children play the violent video games, they're drilling, drilling, drilling--not two times a year--every night, to kill every living creature in front of you, until you run out of targets or you run out of bullets...we're reasonably confident that in Pearl, Mississippi, and in Paducah, Kentucky, and in Jonesboro, Arkansas, these juvenile, adolescent killers set out to shoot just one person: usually their girlfriend,... maybe a teacher. But, then, they kept on going! And, they gunned down every living creature in front of them, until they ran out of targets or ran out of bullets! ... [A]fterwards, the police asked them.... "Okay. You shot the person you were mad at. Why did you shoot all these others? Some of 'em were your friends!'" And the kids don't know." [6] Grossman argues that not only do such games train children to "to kill every living creature in front of you" but, as with real training simulators, the children become excellent shots.

4. Simulation and metaphorisation

'Spatial navigation' on the desktop is achieved by the utilisation of streams of perspectivally rendered images which correspond to the movement of an avatar in a virtual world, combined with arbitrary combinations of mouse movements and keystrokes which correspond to movement on several degrees of freedom (DOF). The notion of 'navigation' in a highly metaphorical 'space' of data is several degrees more abstracted. At this point the notion of 'navigation' is so highly metaphorised that a substantial amount of cultural background is necessary to make the use of the term comprehensible. Web 'navigation' like many computer applications, leverages and metaphorises human skills in spatial location and spatial navigation to facilitate information searches. The degree of literalness of simulation depends substantially upon the precision with which bodily behaviors germane to that task in the real world can be accommodated and measured in a simulator environment. Commercial simulators and many interactive artworks make the assumption that a close and accurate accommodation of bodily behaviors results in a more persuasive simulation. Legible City, Jeffrey Shaws' paradigmatic immersive artwork of 1988, provides a good case example. The interface for Legible City was a stationery bicycle, instrumented so that the speed of pedaling and angle of steering could be extracted. These data directly drove the projection of street-like imagery on a large projection screen in front of the user. The effect was a fairly complete and persuasive simulation of riding a bicycle through a city.

If you play Quake on standard desktop PC, there is no gun sized, gun shaped input device. There is a QWERTY keyboard. A keyboard is not physically like a gun. This is the (naive) argument often made to adduce that first person shooters do not induce violent behavior. But the user's relation to the system is not that simplistic. The user of a first person shooter sees the front end of a weapon on the screen. She can point that weapon in various directions. She can press a key with an index finger to see and hear a plume of fire emanate from the weapon and incinerate some alien beast who writhes in agony in a rewarding fashion before collapsing into a steaming heap.

Many "mouse/keyboard" games can be played with joysticks,which are essentially a pistol grip complete with a trigger. More recent joystick peripherals provides force feedback, the user feels a recoil jolt in the hand when the trigger is pulled. And of course more elaborate arcade game interfaces simulate such effects more completely. Janet Murray notes: "The most compelling aspect of the fighting game is the tight visceral match between the game controller and the screen action. A palpable click on the mouse or joystick results in an explosion. It requires very little imaginative effort to enter such a world because the sense of agency is so direct."[7] This statement demonstrates the way that conventional critiques of representation are rendered inadequate in this fusing of bodily action and real time effect in modeled 3D worlds. The weapon is no longer just a picture. The representation is controlled, driven. In this space between mere pictures and the 'real world' the embodied aspects of simulation influence representation in real time.

5. After 'perception': active sensing.

In the postwar years, theories of visual perception based in gestalt psychology, by such authors as Rudolf Arnheim, Ernst Gombrich, and R L Grergory, had a significant effect in art theory and criticism [8]. From a contemporary point of view, this work, especially when it considered art, was characterised by a conception of vision and visual perception as a one way process of information inflow, through the eyes into the brain. This conception of the detached observer eye, the disembodied mind-like eye, the eye as extension of mind, is dualist and objectivist. The shortcoming of such approaches is that they disregard the dynamic perceptuo-motoric nature of visual learning. In a classic experiment by Held and Hein, a group of kittens were reared in total darkness. The kittens were fitted in a gantry arrangement with two baskets. One basket had holes for the legs such that the physical movements of one kitten would drive both animals through roughly similar spatial experiences. In each case, after a few weeks the kitten which walked and could associate visual information with its own physical movement developed effective vision. The rider in the basket remained functionally blind. [9]

We could never interpret an image of a domestic space, had we not actually moved about in such spaces. Physical experience does not simply disambiguate, it is the key by which images are understood. A baby learns about its visual system via physical exploration and hand-eye experiments. Such experiments might be said to 'calibrate' the babies visual system but in fact they build it, they build a correlation between visual signals and a kinesthetic/tactile nature of the world. Vision is remote sensing but it is grounded on touch and proprioception, on reaching and grasping, stumbling and falling.

In interactive media a user is not simply exposed to images which may contain representation of things and actions. The user is trained in the enaction of behaviors in response to images, and images appear in response to behaviors, in the same way that a pilot is trained in a flight simulator. Passive observation may be shown to have some effect on the beliefs or even the actions of an observer, but an enacted training regime must be a more powerful technique. So critiques of representation derived from painting, photography, film and video are inadequate for discussing the power of interactive experience.

Much debate has occurred on the correlation between pornographic images and sex-crime. Conversations about representations of violence typically conflate movies and computer games, as if they were in the same category. Whatever the power of images, interactive media is more. Not 'just a picture', it is an interactive picture which responds to my actions. Our analysis of interactive media must therefore go beyond theories of representation in images. The image is just the target, the surface. The interactive image cannot be spoken of in the terms of traditional passive images because it is procedural. The content is as much in the routine which runs the image as it is in the image itself. Interactive applications are not pictures, they are machines which generate pictures.

In the theoretical and aesthetic analysis of interactive media, the enactive dimension remains poorly elucidated. In the area of digital media practice which is attached to the visual arts, theories of visual representation have been a powerful critical force. But, as I have been arguing, an interactive 'representation' is more than a representation. In desktop computer based interactives, there are additional levels of metaphorisation (ie of mouse clicks for body movement) but 'embodied interactives' are in a kinesthetically different and more literal territory. While a mouse-click on an html document may 'represent', in a manner similar to a picture, the action of turning a page, it is still in the realm of the symbolic. But actually lifting ones leg and stomping down on a face moves the action several steps more literal. It was not a flesh and blood face, nor a rubber model of a face, but, on the continuum, a responding photo-real interactive image is definitely more 'real' than a block of text or a diagrammatic pistol-shooting target.

When soldiers shoot at targets shaped like people this trains them to shoot real people. When pilots work in flight simulators, the skills they develop transfer to the real world. When children play ‘first person shooters’ they develop skills of marksmanship. So we must accept that such interactive simulation has the potential to build behaviors which can exist without or separate from, and possibly contrary to, rational argument or ideology.

6. Serious play.

My conclusion is that the objection concerning acting and serial killing does not significantly destabilise my basic argument. It may be proposed that engagement with interactive work, or gaming for that matter, is “just a game”, its ‘play’ and therefore doesn’t matter. This turn of phrase obscures the fact that the truth is the opposite: ‘play’ is a powerful training tool. To return to the key question: what behaviors do these games train? (For certainly they do train. Some game production companies are known to implement so-called ‘reinforcement schedules’ based on the long standing psychological knowledge that intermittent reinforcement is a more effective training technique that consistent reinforcement.) In the case of Quake, a simplistic response would be that they improve hand eye coordination at the computer console. This would be true to their cybernetic human-machine interaction roots. Yet this would be to ignore any interaction between the physical action and the imagery and narrative of the game.

We know that simulators do train effectively. Skills learned in simulation are elicited in the world. A pilot responds, 'instinctually' to a situation familiar from training. No reasoning is necessary, indeed, that is exactly the point, this training, like football training is about making responses rapid, reflex. Any Quake playing kid knows how to blow away approaching enemies. Knows in fact, according to the logic of the game, that any approaching stranger *is* an enemy and *must* therefore be blown away immediately. We must assume that these 'learned responses' can also transfer to the real world, if triggered. Such a bodily training may not correlate with any considered political position. Indeed, one might contend that the power of interactive experience is to inculcate behaviors, and these behaviors do not require any ideological correlation. In fact they might work best without such correlation. Such learned behaviors are triggered, without conscious decision-making, when the current context matches the conditions of the training context . (How does the subject know, unconsciously, to produce a specific trained behavior in a specific context? Subtle codes and emotional tenor must play a key roles in such triggering.). There is the possibility that such behaviors might be expressed in situations which resemble the visual context or emotional tenor of the gameplay. Which is to say, games and interactive media in general can be powerful inculcators of behaviors, and these learned behaviors can be expressed outside the realm of the game. And if this is true, then it is hard to escape the conclusion that, for instance, first person shooters actively contribute to an increase in gun violence.

References

[1] Foucault, Michel. Discipline and Punish: The Birth of the Prison. Trans. Alan Sheridan. New York: Vintage, 1977, pp. 3-8.

[2] Maus, Marcel. Techniques of the Body. in Economy and Society 2.1, 1973 Ben Brewster Translation . original : Journal de Psychologie Normal et Pathologique, Paris Année xxxii 1935, pp271-93.

[3] Bordieu, Pierre. Outline of a Theory of Practice. New York: Cambridge UP, 1977.

[4] http://www.ict.usc.edu/

[5] Grossman, David. On Killing: The Psychological Cost of Learning to Kill in War and Society. Little, Brown and Co., 1996

[6] Excerpt from Interview with David Grossman: Giving Children the Skill and the Will To Kill. http://members.tripod.com/~american_almanac/grossint.htm.

[7] Murray, Janet. Hamlet on the Holodeck, the future of narrative in cyberspace. New York, free Press, 1997. Reprinted by MIT, 1999, pg146

[8] Arnheim, Rudolf. Art and Visual Perception: A Psychology of the Creative Eye. (Berkeley: University of California Press, 1954); Arnheim, Rudolf. Toward a Psychology of Art. (Berkeley: University of California Press, 1966). Gombrich, E.H. Art and Illusion: A Study in the Psychology of Pictorial Representation, Princeton 1961, Gregory, R L. Eye and Brain: The Psychology of Seeing. Princeton 1966

[9] Held, R, and A Hein, Adaption of disarranged hand-eye coordination contingent upon re-afferent stimulation. Perceptual-Motor Skills 8. 1958 pp87-97. Also cited in RL Gregory, Eye and Brain, Fourth edition. Princeton University Press 1990, p219

Assessing Truly Episodic Memory with a Virtual Environment: Effect of Aging, Encoding and Sensory-motor Implication

Plancher, G.
gaen.plancher@univ-paris5.fr

Nicolas, S.
serge.nicolas@univ-paris5.fr

Piolino, P.
pascale.piolino@univ-paris5.fr

Laboratory of Psychology and Cognitive Neurosciences, CNRS, University Paris Descartes.
Research group Memory and Learning
71 avenue Edouard Vaillant, 92774 Boulogne-Billancourt, France

Abstract

Episodic memory is rarely studied in a frame closed to its complete definition. We studied episodic memory for young and older adults in active and passive exploration and in intentional versus incidental encoding in a virtual environment. The active participants drove a car through a virtual town whereas passive participants were only passengers. After the exploration, subjects performed memory free recall test which assessed all components of episodic memory; the verbal components: the what, the when, the where, the details; the visuo-spatial component; and they took a recognition test. Our findings showed no difference between age groups in incidental encoding regardless of the components. By contrast, young adults were better in intentional encoding for target, verbal-spatial and for temporal information; but not for the visuo-spatial score. Old adults showed better recognitions than young. In addition, we found no effect of enactment on episodic memory. Virtual reality appears as an interesting tool for study episodic memory.

1. Introduction

According to Tulving [10], episodic memory allows conscious recollection of personal events, together with their phenomenological and spatio-temporal encoding contexts. The constructive memory framework postulates that the features of an episode are linked together at encoding to form a coherent representation. The recollection of one element of the episode allows the recollection of the other elements [5].

Yet, the majority of episodic memory studies are far from what we experience in daily life and far from its definition [10]. In order to investigate all components of episodic memory a context where we can assess the memories of *what*, *when*, *where* and *details* of an event is needed. Virtual reality (VR) can provide a fully controlled experimental situation in a rich context and it allows active interactions between the subject and the stimuli.

Difficulties with episodic memory are among the most frequent cognitive deficits in normal aging. Aging deficit is more important for the memory of the spatio-temporal context than for the memory of factual information. For example, older adults are less likely to correctly remember contextual features of events, such as their colour or location [3]. The capacity for retrieve an event can depend on the quality and on the strength of associations created between central information and contextual features [7]. The motor activity and the intentional encoding can reinforce association between features of an event.

The fact to be active rather than passive in an event enriched the context of episodic memory. It has been shown that performance in episodic memory is enhanced by this activity [11]. This effect is usually attributed to the good item-specific information provided by enactment. It was shown also for spatial memory in virtual environment [1,2]. Moreover, it was observed that older adults have a deficit for spatial memory in virtual environments [6]. Active compare to passive exploration in virtual environment would enhance episodic memory of participants. Another factor which contributes to a richer memory trace is the intentionality of learning. Participants show better memory performance when they learn intentionally the items, i.e. they try to memorize it, than when they learn the items incidentally, i.e. they are not informed that their memory will be tested [8]. Intentionality should enhance episodic memory.

In the present study we were interested on the memory of different components of episodic memory for young and older adults in a virtual environment. We compared their performances in active and passive exploration and in intentional versus incidental encoding.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

2. Methods

2.1. Participants

113 undergraduate psychology students from Paris Descartes University (67 female and 33 male; M_{age} = 21.57; *SD* = 2.99) and 45 older adults participated voluntarily (M_{age}=59.42; SD=9.85). They were randomly assigned to one of the four conditions resulting of the crossing between active/passive sensory-motor implication and intentional/incidental encoding. Inclusion of subjects was based on absence of neurological or psychiatric medical history and signs of depression. No medication known to impair memory was allowed. The groups did not differ on verbal abilities according to the 44-item Mill Hill test. Old adults shown a normal score at the MMSE and had a socio-cultural level equal to young adults.

2.2. Materials

2.2.1. The Virtual Apparatus

The virtual apparatus was composed of a computer generated 3-D model of a created environment. The environment was built with Virtools Dev 3.0 software (www.virtools.com). The environment was run on a PC laptop computer and explored using a steering wheel, a gas pedal and a brake pedal. It was projected with a video projector onto a screen (85 cm high and 110 cm long).

2.2.2. The virtual environment (VE)

A town based on photos of Paris, France, was built. In the town, there was one possible route composed by nine turns. Nine *specific areas* with their *context* composed the town's elements. The interconnected *specific areas* were, in this order: the area of tall buildings, the shops, the roadblocks, the town hall, the restaurants, the car accident, the train station, the arcade buildings, the old red buildings and the park. One of these areas is a car accident scene; in this specific event, two cars crashed into one another, a horn was heard and black smoke appeared. Buildings connected each area with another. People, garbage containers, barriers, trees, billboards and motionless cars were the *context* of the town. Each area was composed of some contextual elements. For example, in front the train station: a car accident occurred, a woman walked, a billboard and trees were presented (see Figure 1.)

Figure 1. Example of a virtual town's view

2.3. Procedure

Participants were tested individually. They were seated on a comfortable chair. The VE was projected on the wall, 150cm ahead of subjects. In order to match the visual stimulation between active and passive participants, we recorded the navigation of all active participants. Each passive participant saw the record of one active participant.

If the subjects were in an incidental encoding, they were told to drive through the virtual town and pay attention to the town. If the participants were in an intentional encoding, we asked participants to try to remember the most elements they saw in the town and the itinerary of the driver in order to recall them at the end of the presentation.

2.3.1 Exploration of the virtual environment

Active condition

The participants manipulated a steering wheel, a gas pedal and a brake pedal to move around in the virtual town. Before the immersion within the town, the participants trained in an empty environment until they were comfortable. We notified the subjects that they did not have to drive too fast. Then, the active participants were immerged in the virtual town. Only one trip was possible, the participants could not stop and could not turn back. Navigation stopped at the end of the route, approximately two minutes after the start.

Passive condition

The participants were immerged directly in the virtual town. We told subjects they were sitting on the passenger side. The passive participants saw the recorded route of the active participants and in order to control the attention in our experimental situations; we told passive subjects to pay attention to the driving and to the itinerary of the driver.

2.3.2 Memory tests

The free recall test

What recall

This recall was scored out of a possible 31 elements. Participants had "to recall all the elements presented in the town". The maximum was 9 central elements (e.g. "the train station") and 22 minor elements (e.g. "the girl").

Where recall

This recall was divided in two different parts: the verbal *where* recall (*where 1*) and the visuo-spatial *where* recall (*where 2*). For the verbal *where*, the subjects had to remember spatial information concerning the element, if it was "in front of us", "on their left" or "on their right". This recall was scored out of a possible 31 verbal *where*. For the visuo-spatial *where*, the subject had to draw the map of the town and localise the elements on the map. The map score was computed with the number of correct turns, the maximum score was 9 turns. Participants had also to recall the locations of elements on a correct map; the maximum score was 31 correct locations. Thus, the maximum of visuo-spatial *where* score was 40 correct recalls.

When recall

31 temporal recalls were possible. The participants had to recall when the element occurred, if it was "at the beginning", "at the middle" or "at the end of the town".

Details recall

22 *details* recalls were possible. The subjects had to recall the details of the element they remembered (e.g. one of the cars was blue).

The recognition test

A recognition test was presented after the recall. The subjects had to choose the item presented in the town among three different items. It was composed of 10 questions concerning the elements and their locations in the town. For example, "who was presented in front of the accident?"

3. Results

Analyses of variance (ANOVA) were conducted on the data recorded from the different components of the episodic memory: the *what*, verbal *where*, visuo-spatial *where*, *when* and *details* and on the correct recognitions. The encoding, the sensory-motor implication and the age were included in the ANOVA as between factors.

A main effect of the encoding was found, the intentional encoding compared to the incidental encoding leads to a better recall of *what* (F (1,152) = 8.04; $p < 0.01$), of verbal *where* (F (1,152) = 9.29; $p < 0.01$), of visuo-spatial *where* (F (1,152) = 9.93; $p < 0.001$), of *when* (F (1,152) = 9.31; $p < 0.01$), of the *details* (F (1,152) = 4.35; $p < 0.05$). Learning information intentionally leads to better memory of all components of episodic memory.

Second, the sensory-motor implication had no effect neither on the *what* (F (1,152) < 1), the verbal *where* (F (1,152) = 1.24; n.s.), the visuo-spatial *where* (F (1,152) = 1.12; n.s.) the *when* (F (1,152) = 1.77, n.s.), the *details* (F (1,152) < 1) recall scores nor on the recognition test (F (1,152) = 1, n.s.). Activity did not enhance the different components of episodic memory in a VE.

Moreover, we found a main effect of the age, young adult were better than old adults on verbal *where* recall (F (1,152) = 65.94, $p < 0.001$), on visuo-spatial *where* recall (F (1,152) = 4.8, $p < 0.05$), on *when* recall (F (1,152) = 14.6, $p < 0.001$), on *details* recall (F (1,152) = 7.9, $p < 0.01$). No effect of the age was found on the *what* recall (F (1,152) = 2.28, n.s.). Though, old adults showed a higher score on the recognition test (F (1,152) = 11.88, $p < 0.001$).

In addition, we found interactions between the encoding and the age (See Figure 2.). The young adults were better only in intentional encoding for the *what* (F (1,112) = 6.81, $p < 0.01$), the verbal *where* (F (1,112) = 13.59, $p < 0.001$), the *when* (F (1,112) = 13.08, $p < 0.001$). Moreover, old adults were better than young only in incidental encoding for the recognition score (F(1,112) = 4.91, $p < 0.05$). No interaction was observed for the visuo-spatial *where* score (F (1,112) = 2.65, n.s.) and for the *details* score (F(1,112) < 1).

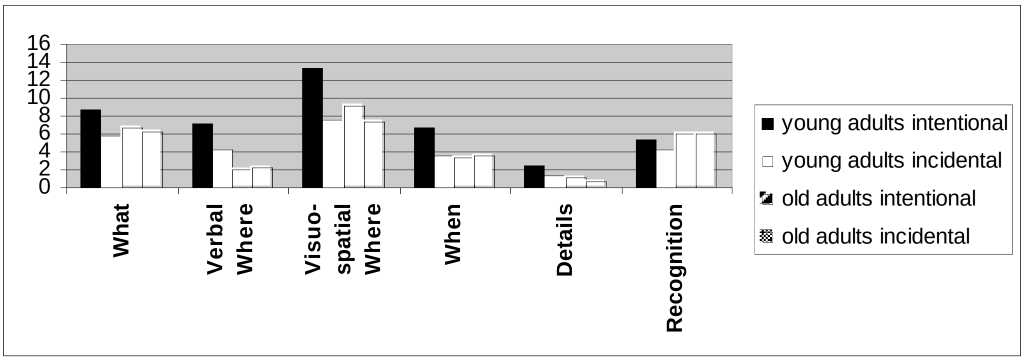


Figure 2. Recall of various episodic memory components according to the age and the encoding.

4. Discussion

It was the aim of the current study to examine the effects of aging, encoding and sensory-motor implication on various components of episodic memory. The results clearly indicate age-related decline on memory of the spatio-temporal context and of details unlike factual recall. These findings are in agreement with previous results showing that contextual memory decline with aging [3].

However an interaction was found between the age and the encoding: no difference appears between ages in incidental encoding. By contrast young adults were better in intentional encoding for target (*what*) information, for verbal spatial information (*where 1*) and for temporal information (*when*) than old adults. In addition what was surprising is the superiority of older adults compare to young in incidental encoding for correct recognitions. An explanation for this effect could be that this test was especially about the accident and it is possible that older adults were more attentive to this emotional event. Thus they recognized stimuli better than young adults who were more used to with video games and their emotional stimulations. But old adults were better for recognitions only in incidental encoding. The results showing that older adults did not benefit from intentional encoding suggest that age-related deficits in memory may come from the incapacity for old adult to develop encoding strategies. However, this interaction was not observed for the visuo-spatial score (*where 2*), i.e. the old adults benefit of intentional encoding for this score. So this finding leads to believe that intentional encoding is most likely to be effective for older adults when information is visuo-spatial rather than verbal. Therefore, for young adults encoding was advantageous for all scores even for visuo-spatial score which is more implicit and automatic. These data confirms previous studies showing that spatial memory is function to intention to learn for young and old adults [9]. Though, this result is evidence against the total automaticity of the encoding of spatial location information [4].

Moreover, we observed no effect of the sensory-motor implication on the recall of elements and on their spatio-temporal context. These results are paradoxical with studies using the enactment paradigm showing that action enhances memory [1,2]. But in previous studies, the motor activity was composed of more movements than in our study. In our study, visuo-motor interaction and motor control were not that such as important. We can conclude that memory of these components do not require visuo-motor coupling in a VE, passive exploration is sufficient to this cognitive ability. We propose that no effect of active exploration on episodic memory occurred on episodic verbal scores because the subject verbalized what he perceived and then recalled verbal information he inferred from perception rather than the perceived information. In addition, we supposed that no effect was observed on visuo-spatial score because our motor activity is very weak, no comparable with previous studies showing a sensory-motor effect. It should be interesting to study the effect of enaction on episodic memory with a bigger action. Last, VE appears to be an appropriate tool to study evolution of truly episodic memory for young and older adults.

References

[1] Brooks, B.M., Attree, E.A., Rose, F.D., et al. (1999). “The specificity of memory enhancement during interaction with a virtual environment”. Memory pp.65-78.

[2] Carassa, A. Geminiani, G., Morganti, F., et al. (2002). “Active and passive spatial learning in a complex virtual environment: The effect of efficient exploration”. Cognitive Processing, pp.65-81.

[3] Chalfonte, B. L. and M. K. Johnson (1996). "Feature memory and binding in young and older adults." Memory & Cognition, pp. 403–416.

[4] Hasher, L. and Zacks, R.T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General, pp.356-388.

[5] Kessels, R., Hobbel, D., & Postma, A. (2007). “Aging, context memory and binding: a comparison of “what, where and when” in young and older adults”. International Journal of Neuroscience, pp. 795-810.

[6] Kessels, R.P.C., Te Boekhorst, S. and Postma, A. (2005). "The contribution of implicit and explicit memory to the effects of errorless learning: A comparison between young and older adults." Journal of the International Neuropsychological Society, pp.144-151.

[7] Lekeu, F., Van der Linden, M., Moonen, G. and Salmonet, E. (2002). "Exploring the effect of action familiarity on SPTs recall performance in Alzheimer’s disease." Journal of Clinical and Experimental Neuropsychology, pp.1057-1069.

[8] Neill, W.T., Beck, J.L., Bottalico, K.S., et al. (1990). “Effects of intentional versus incidental learning on explicit and implicit tests of memory”. Journal of Experimental Psychology: Learning, Memory and Cognition, pp.457-463.

[9] Park, D.C., Puglisi, J.T. et al. (1982). "Spatial memory in older adults: effects of intentionality." Journal of Gerontology, pp.330-335.

[10] Tulving, E. (2002). “Episodic memory: from mind to brain”. Annual Review of Psychology, pp. 1-25.

[11] Zimmer, H. D., Cohen, R. L., Guynn, M. J., et al. (2001). Memory for action. New York: Oxford University Press.

Towards a Flexible Real-time Gesture Recognition System for Virtual Environment Control

Otniel Portillo-Rodriguez* Oscar O. Sandoval-González* Haakon Faste*
Jesús A.Velázquez Lechuga*† Gerardo A. Saucedo Basilio*† Elvira Todaro*
Carlo Alberto Avizzano* Massimo Bergamasco*
(*)PERCRO Laboratory, Scuola Superiore Sant'Anna, Italia
(†) Instituto Tecnológico de Monterrey Campus Toluca, Mexico*
E-mail: otniel@sssup.it, bergamasco@sssup.it

Abstract

The research outlined in this paper integrates diverse technologies in the development of a flexible gesture recognition system allowing users to record human movement, define gestures within the recorded data and test the recognition of the recorded gesture set. Such research has potential applications in many fields including communications, dance, education and HCI.

The system is designed to be easily configurable for a variety of different tasks and applications, and integrates technologies including VICON motion tracking system, MatLab, Neural networks and XVR. The system was tested using seven gestures focusing on arm and hand movement with 98% accuracy.

1. Introduction

Optical gesture recognition is an important area of research with a wide array of potential applications in many fields including dance, sport, rehabilitation, communication, motor learning, education, and HCI. The goal of this research was to develop a flexible system allowing body gestures to be recorded, identified and archived so that subsequent gestures could be recognized by the system with high accuracy.

There are many approaches to the design of such systems that use a variety of methodologies to model, analyze, and recognize gestures [1]. Machine Learning and Pattern Recognition software is often used to reconstruct various motions from video sequences [2] or motion capture systems [3]. Such cues can then be used to drive interactive systems. Camurri et al. [4] have used the Laban Effort movement qualities [5] to guide music synthesis, Woo [6] extracted the speed and trajectory of limb movements around the body, and Moeslund [7] used different static human body gestures (poses) to drive an interactive dance environment.

The gesture analysis process used in this research follows several inter-related steps: kinematic model construction and gesture recording using a VICON 3D motion capture system, gesture analysis, data processing, gesture recognition via Artificial Neural Networks, and gesture model testing. A detailed technical overview of these steps is described in the following sections.

2. Experimental Setup

2.1. Capturing Environment

The VICON system's cameras are placed around a close area for create a capture volume, as shown in figure 1. The system is composed of 8 cameras that track the positions of retroreflective markers attached to the user performing the movement.

Figure 1. The capturing environment and a marker.

The position of the markers was chosen in order to acquire only movements of the arms in a way that they can be easily affixed to the user and recorded. Fifteen markers were used, as shown in figure 2.

A model template was created with the markers using VICON's Nexus Software. This template has nine segments, between each of which is a ball joint as shown in figure 3. The model allows the velocities, angles and positions of gesture performance to be seen, helping the operator understand which information is relevant to defining the movement.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

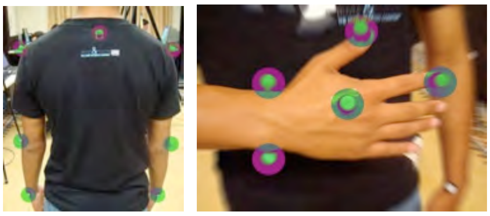


Figure 2. Marker placement on the arms and hands

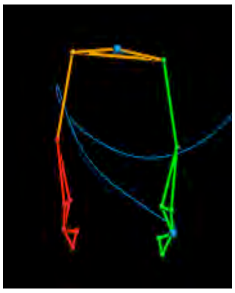


Figure 3. Kinematic structure in Nexus

To acquire the markers' positions in a way that they could be used with other applications, an interface between VICON's Nexus Software and XVR, an online virtual environment development platform [8][1].

2.2. Captured Gestures

The movements for the first test of the project are classified into symmetric gestures (tracing symmetric figures in space with two hands) and asymmetric gestures (ex. Police officer controlling traffic). Two symmetric gestures were recorded in addition to five asymmetric gestures, as shown in figures 4 and 5. All movements have an initial and final position and it is fixed.

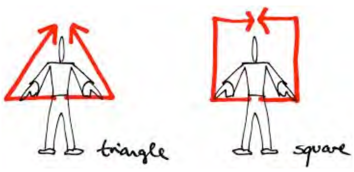


Figure 4. Symmetric gestures

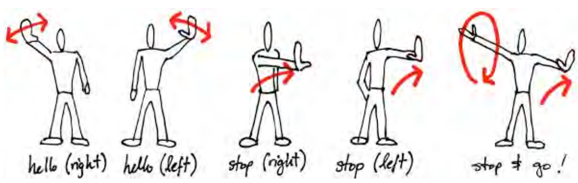


Figure 5. Asymmetric gestures

3. Description of gesture modeling

After analyzing the behavior of the movements, it was determined that the only attribute which does not change significantly depending on the user's position, placement or body size are the angles of the kinematic structure. With this in mind, the values of the angles between the finger/thumb, palm/forearm, forearm/arm and arm/shoulder of each arm were calculated based on the markers' positions shown in figure 6.

Figure 6. Obtained angles used for gesture recognition

The plotted angles revealed that some values don't seem to be needed because they do not belong to the movement itself. For this reason, a "chopper" was developed. The idea behind the chopper is to use the angle velocities and palm velocities in order to know when the desired movements start and finish. These values are visualized for a sample gesture in figures 7 and 8.

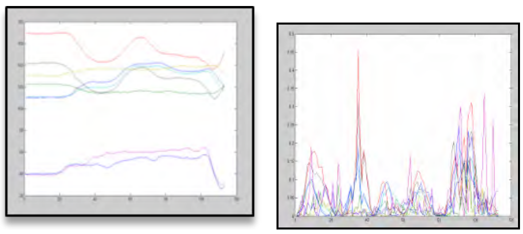


Figure 7. Angles and angular velocities over time

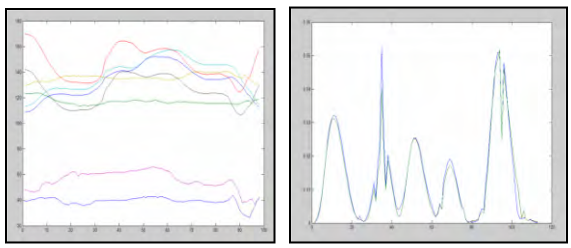


Figure 8. Isolated angles and palm velocities over time.

4. Software and computing architecture

The overall sequence of the system architecture is shown in figure 9, beginning with the user performing movements to record the markers placed on the body.

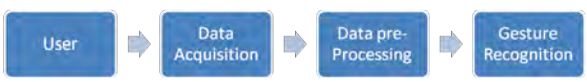


Figure 9. Information flow

The software architecture for the system includes the elements shown in Figure 10. The software used was VICON's Nexus control program, XVR and Matlab Neural Networks toolbox version 5.0.1.

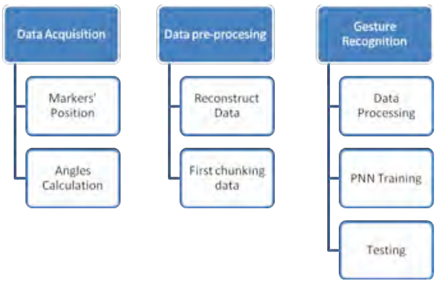


Figure 10. Software architecture

4.1. Data Acquisition

All the markers' positions are obtained via the VICON system using VICON's Nexus Software, which also constructed the upper body kinematics. The data acquisition module then interfaces Nexus to XVR by means of a dynamic link library (DLL). In XVR the positions (x, y, z) of each markers (acquired every 20 milliseconds for 5 seconds) are stored into an array. Then, the array is used to calculate all the angles saved in a TXT file.

4.2. Data pre-Processing

For each TXT file, a series of manipulations are performed on the data. The first algorithm corrects the angles in the case of a lost marker. Next, another algorithm calculates the velocities of the palms and the angles (with basic mathematical methods). This is used to determine where each movement starts and finishes, and to filter the data from the gesture. A subsequent algorithm uses these values to chunk the movement, based on when the first velocity (either angular or of the palms) changes from zero (start) and returns to zero and remains that way for some time (end).

4.3. Data Processing for Gesture Recognition

Seven basic gestures outlined in section 2.2 were used. Each gesture was described by 8 angles, and 25 different trials of every gesture were taken.

In order to feed the data correctly to the neural networks used to recognize the gestures, it was necessary to create two matrixes: one composed of the training trial gestures, and the other composed of the test trial gestures. In each matrix, every column represents one gesture. 15 of the 25 trials of each gesture were used as training gestures, and the other 10 were used as test gestures.

Each gesture is preprocessed as described in section 4.2. The 8 angles of each gesture are chunked into 20 parts, creating a vector with 160 values.

4.4. Gesture Recognition

Two neural networks were used to recognize the gestures: Feed Forward Neural Network (FFNN) and Probabilistic Neural Network (PNN) [9]. The FFNN is normally used as a general approximate function, so in this case it was necessary to test whether or not it could distinguish gestures presented as input. The architecture of the FFNN was 160 input neurons, a variable number of hidden neurons, and 7 output neurons.

In the PNN when an input is presented, the first layer computes distances from the input vector to the training input vectors and produces a vector whose elements indicate how close the input is to the training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a competitive transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 for that class and a 0 for the other classes.

5. Early experiments & evaluation of results

The training matrix and its corresponding class gestures were fed to the FFNN and PNN in order to test them. All of the tests were performed in Matlab version 7.3 on a PC with a 1.8 Ghz Centrino microprocessor and one gigabyte of memory.

The first experiment consisted of checking the performance of the FFNN to recognize gestures while varying the number of neurons in its hidden layer, while keeping the number of trajectory points of every angle, the chunking algorithm, etc., invariant. Table I shows the average results obtained for 100 different instances of the FFNN. A linear relationship does not exist between the number of neurons and the percent error, and the necessary time to create and train the neural network was at least 26 seconds. Furthermore, it is important to notice that each instance of this neural network always provides different results. The best result was obtained with an FFNN of 12 hidden neurons (98% of accuracy).

The second test was over the PNN. The architecture of this neural network is fixed, and depends only on the number of elements in the input vector and the number of classes to identify. Furthermore, each instance of this network created always provided the same results, although the creation and evaluation times may be different [9]. Table II shows the average results obtained for 100 instances of the PNN network. In general this network was fast to create and trained with 98% accuracy.

Table I. Results Obtained for the FFNN

Hidden Neurons	Time evaluation of the test matrix (seconds)	Time Creating and Training (seconds)	Percent Error
7	0.016	350.531	31.4286
8	0.016	501.359	18.5714
9	0.015	593.031	41.4286
10	0.016	26.359	2.8571
11	0.016	940.781	22.8571
12	0.016	32.672	1.4136
13	0.015	42.11.00	2.8571
14	0.016	37.641	5.7143
15	0.017	358.453	11.4286

Table II. Results Obtained for the PNN

Time evaluation of the test matrix (seconds)	Time Creating and Training (seconds)	Percent Error
0.0395	0.0472	1.4286

The repeatability of the results in each instance of the PNN suggest that this neural network is the best option for the real time implementation of the Gesture Recognition System for Virtual Environment Control.

7. Future Work

There are certain limitations of the VICON system and experimental setup that should be overcome. For example, the current system involves manually placing markers on each user, which may not be practical in learning environments. Creating a simple suit consisting of gloves and Velcro straps, for example, could be an easy solution to address this issue.

Because strangers to the system will not know how to perform the given gestures, some degree of visual instruction should be integrated. This may include interactive video allowing a user to watch a video clip and perform the demonstrated gesture, or an avatar/teacher to demonstrate the ideal movement. In the latter case, the use of realtime stereoscopic visualization would be a powerful addition, as would “mirror learning” in which the student sees his or her own kinematic-skeleton reflected in a virtual “mirror,” perhaps superimposed with that of the correct gesture.

To date the system provides accurate gesture recognition for gestures with a clear starting and stopping position, but real gestures are often part of a fluid sequence. Improving the system to account for continuous motion is a challenging area of research.

To enhance motor learning, elements providing haptic feedback to correct faulty posture could be incorporated into the system. Vibrotactile feedback suits, for example, have been shown to be effective aids to accelerated human motor learning [10][11].

8. Acknowledgements

In addition to the Enactive Network of Excellence, special thanks to the SKILLS IP consortium, an EU funded project dealing with the acquisition, interpretation, storing and transfer of human skill by means of multimodal interfaces, robotics and virtual environment technologies.

References

[1] Ying Wu, Thomas S. Huang, “Vision-Based Gesture Recognition: A Review”, Springer Berlin/Heidelberg, 1999

[2] K. Pullen and C. Bregler, “Motion capture assisted animation: texturing and synthesis,” In SIGGRAPH, pp. 501-508, 2002.

[3] L. Ikemoto and D. A. Forsyth, “Enriching a motion collection by transplanting limbs,” In Proceedings of the 2004 ACM SIGGRAPH/Eurographics symposium on Computer animation, pp. 99-108, 2004.

[4] Camurri, A., Hashimoto, S., Ricchetti, M., Ricci, A., Suzuki, K., Trocca, R. and Volpe, G., “EyesWeb: Toward Gesture and effect Recognition in Interactive Dance and Music Systems,” *Computer Music Journal. 24 (1)*, 57-69, 2000

[5] Moore, C.-L., Yamamoto, K., *Beyond Words: Movement Observation and Analysis*. Gordon and Breach Science Publishers, New York, 1988.

[6] Woo, W., Kim, N., Wong, K., and Tadenuma, M., “Sketch on Dynamic Gesture Tracking and Analysis Exploiting Vision-based 3D Interface,” *Proc. SPIE PW-EI-VCIP'01, vol. 4310,* pp. 656-666, 2000

[7] Moeslund, T. B., and Granum, E. A., “Survey of Computer Vision-based Human Motion Capture,” *Computer Vision and Image Understanding 81 (3),* pp. 231-268, 2001

[8] VRMedia – eXtremeVR: virtual reality on the web, viewed 19 July 2007, http://www.vrmedia.it

[9] Bishop, Christopher M., “Pattern Recognition and Machine Learning.” ISBN 780387310732. Springer, 2006

[10] Avizzano, C.A., M. Bergamasco, “Haptic Interfaces: a New Interaction Paradigm”, Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, 1999

[11] Lieberman, J., and Breazeal, C., “Development of a Wearable Vibrotactile Feedback Suit for Accelerated Human Motor Learning”, IEEE International Conference on Robotics and Automation, Roma, Italy, 2007

Enactive Navigation in Virtual Environments: Evaluation of the Walking PAD

Paul Richard* Laroussi Bouguila† Michèle Courant† Beat Hirsbrunner†
*Laboratoire d'Ingénierie des Systèmes Automatisés – University of Angers, France
†Department of Informatics (PAI Group), University of Fribourg, Switzerland
E-mail: richard@istia.univ-angers.fr, michel.courant@unifr.ch, Laroussi.bouguila@unifr.ch

Abstract

This paper presents a human performance evaluation of a low-cost enactive locomotion interface, the walking-PAD, that provides users with the ability to engage in a life-like walking experience in virtual environments (VEs) by stepping in place. Stepping actions are performed on top of a platform with embedded grid of switch sensors that detect footfalls pressure. Based on data received from sensors, the system computes different variables that represent user's walking behavior such as walking direction and walking speed. Twelve human subjects were instructed to reach the exit of a virtual labyrinth as quickly as possible and memorize as many information as they can. Two navigation techniques were compared: a mouse-based technique and the walking-PAD technique. Results revealed that more information was memorized when using the walking PAD.

1. Introduction

How users move from one virtual place to another within limited tracked space is one of the most difficult and persistent problems for immersive virtual environments (VEs). Flying by pushing a button or joystick has none of the naturalness of really walking because it doesn't make users tired or stimulate users proprioceptive and vestibular systems. Real walking overcomes those problems, but introduces the problem of how to move in VEs that is larger than the area covered by the tracking system. Iwata [6] and Darken [4] proposed two different locomotion interfaces with similar principle of omni-directional treadmill systems, which can cancel users' displacement and keep them located at the same place while being able to walk into any direction. Slater [13], Templeman [14] and Usoh [15] adopted a simpler approach that eliminates the need of moving platform to cancel user displacement.

Figure 1. Illustration of the step-in-place technique using the Walking-PAD

They used the step-in-place technique to engage user into a walking experience. Such action keeps a very similar body movement to actual walking behavior but without body propulsion.

Razzaque et al. [10,11] and Bouguila et al. [7] have used the stepping in place with interesting redirection techniques that keep user oriented toward the screen in a non-immersive configuration. Other locomotion systems have been developed to reproduce active walking experience within VE such as bi-pedals interface or linear treadmill [2]. These interfaces do not provide users with the ability to use their active and natural body turning action to change their walking direction. Instead omni-directional locomotion is achieved by using an extra artificial interface to accomplish a rotation task.

This paper presents a human performance evaluation of a low-cost omni-directional enactive locomotion interface that can be integrated into a wide variety of VR systems equipped with any visual display. The interface employs step-in-place technique and a sensitive walking platform to impart users with the ability to freely engage in a life-like walking experience into any direction.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19^{th}-22^{nd}, 2007

The proposed interface was designed to promote the following points:

- Body centered: the locomotive omni-directional actions controlling the navigation are initiated and sustained by the lower part of the body as in real life. This approach will preserve user's natural reflexes and navigational control skills. Moreover, the system lets user's hands free for manual interaction.
- Simplicity: the walking interface is easy to set up, easy to learn, and easy to use, decreasing the mental workload due to the interface.

2. System overview

A human-scale multi-modal VR platform was developed to provide users with the ability to both navigate and manipulate virtual objects. The platform promotes enaction by supporting both multi-sensory (visual, auditory, haptic, and olfactory) modalities and real world based interaction techniques. In particular, haptic interaction is realized using a bimanual human-scale stringed-based haptic interface [12]. Locomotion is achieved using a low-cost easy-to-use enactive interface that permits users to control both their speed and direction of walking in virtual worlds [8].

2.1.The PAD platform

The platform illustrated in figure 2 has a compact size of 45cm x 45cm and weights less than 7kg. A total of 60 iron switch sensors are embedded on a plexiglas surface. The sensors are placed in matrix form to allow locating footfalls and walking direction during stepping actions. The PAD is connected to a PC computer through an NI DAQ board. A second PC computer is used for graphics and receives computed data such as walking speed and direction. Taking into account real-time interaction, the system is set to scan all sensors at a rate of 100Hz. Values that can take each sensor is either "0" or "1", the equivalent of "on' and "off". "0" indicating a free status whereas "1" reflects the presence of the foot on top of the sensor. Values collected from each sensor are denoted as W_n (n=1,2…60) and can be 0 or 1, which represent the sensor weight. To compute the center of gravity (CoG), only activated sensors are taken into account. Therefore, after fetching all sensors values, the CoG is computed based on equation 1.

$$\begin{aligned} P(t) &= (0,0) \qquad if \sum W_{ij} = 0 \\ P(t) &= \sum W_{ij} P_{ij} \Big/ \sum W_{ij} \end{aligned} \tag{1}$$

Where *P(t)* represents the CoG coordinate and P_{ij} represents the coordinate position (x,y) of sensor S_{ij}.

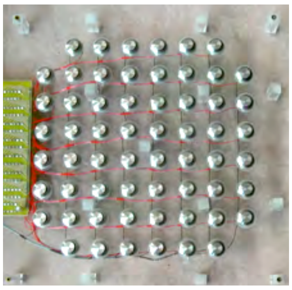


Figure 2. Top view of the Walking-PAD Platform

2.2. Moving direction

Moving direction is obtained by getting the perpendicular axe that passes through the middle of left and right foot falls. As stepping action has certain frequency, it is useful to take the average position during certain interval of time. Figure 3 shows an overview of two different walking directions and their respective CoG plotting.

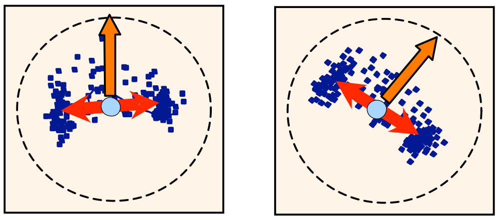


Figure 3. Moving direction based on CoG data

3. Experiment

This section reports an experimental study that aims to compare a mouse-based navigation technique with the walking-PAD navigation technique. Our assumption is that the walking-PAD navigation technique supports enactive knowledge through intuitive movements and involves a better memorization of information from the VE.

3.1. Virtual environment

The experimental VE consists of a labyrinth (Fig. 4). Four amphoras, numerated from 1 to 4 are placed in the labyrinth in order to help the subjects by providing intermediate target. The last amphora is considered as the exit. Five "Star Wars" movies posters and two aubade posters were placed onto the walls of the labyrinth.

3.2. Participants

A total of twelve volunteers' male students from ISTIA (Angers University) participated in the experiment. They were aged from 20 to 26 years old and they never experienced VR before.

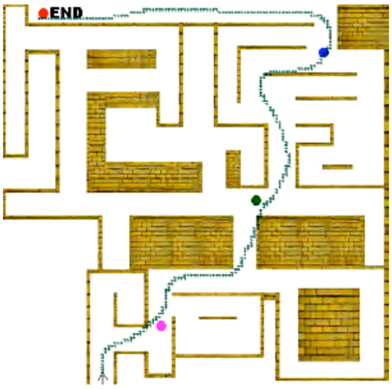


Figure 4. Top view of the labyrinth: a printed path shows a direct way from the entrance to the exit

3.4. Task

The participants were asked to reach as quickly as possible the exit of the virtual labyrinth, while passing through the three intermediate targets (amphoras). They were also asked to memorise as many information as they can from the VE.

3.5. Design

Participants were separated in two groups of six each. The first group (G1) performed the task using the mouse-based navigation technique, while the second group (G2) performed the task using the step-in-place technique.

3.5. Procedure

Before the experiment started, and after looking for 30 seconds to the top view of the labyrinth, each subjects performed the task one time to get acquainted with the navigation technique they were to use in the experiment. All participants were positioned at 2.5 meters from the large screen. Experimental session consisted in three trials. At the end of the experiment, all participants were asked to fill a questionnaire. Questions were about the preferred navigation technique, the number of posters and information on the posters.

Figure 5. View of the entrance of the labyrinth with the first amphora and a "Star Wars" movie poster

4. Results

Results are presented and analysed according to two criteria: task completion time and information recall.

4.1. Completion time

Task completion time was recorded for each single trial. This data was analysed using ANOVA. Results illustrated in Figure 6, revealed a statistical significant time difference between conditions $F(11,1) = 13.34$; $p < 0.005$. The mouse-based navigation technique gave better results. Participants of group G1 have reached the exit of the labyrinth in an average completion time of about 3.6 min. (std: 0.57), while participants of group G2 performed the task in an average of about 5.0 min. (std: 0.60 min.). This result shows that the mouse-based navigation technique was easier to use than the other to navigate in a virtual labyrinth. However, data from the questionnaire reveals that the subjects found the step-in-place technique more intuitive and immersive.

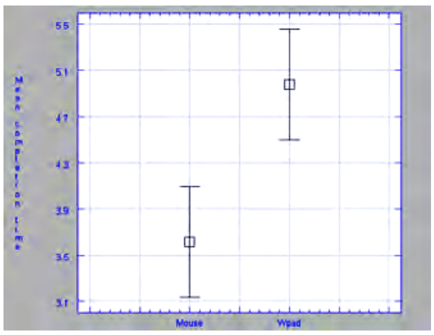


Figure 6. Mean task completion time vs. navigation technique

4.2. Memorisation

The number and type of information memorised were obtained from the questionnaire and analysed using ANOVA. Results illustrated in Figure 7, revealed a significant difference between conditions F (11,1) = 7.50; p <0.005. The step-in-place technique gave better results. Participants of group G1 had an average of 3.33 good results (std: 0.52), while participants of group G2 had an average of 4.33 good results (std: 0.51).

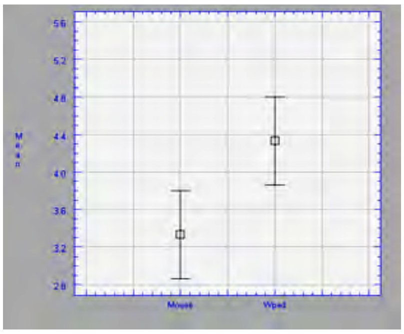


Figure 7. Mean number of information memorized vs. navigation technique

5. Conclusion

This paper presents a human performance evaluation of a low-cost enactive locomotion interface that provides users with the ability to engage in a life-like walking experience in virtual environments (VEs) by stepping in place. A total of twelve volunteers' male students aged from 20 to 26 years were instructed to reach the exit of a virtual labyrinth as quickly as possible and memorize as many information as they can. Two navigation techniques were compared: a mouse-based technique and the walking-PAD step-in-place technique. Results showed that the mouse-based navigation technique was easier and to use and faster than the other. However, subjects found the step-in-place technique more intuitive and immersive. Indeed, this technique gave better results for memorisation of information from the VE.

References

[1] Brogan, D. C., Metoyer, R. A., & Hodgins, J. K. "Dynamically simulated characters in virtual environments", SIGGRAPH7 Visual Proc., pp. 216, 1997.

[2] Couvillion, W. Lopez, R. Ling, J. "The Pressure Mat: A New Device for Traveling Virtual Environments Using Natural Motion," Proc. of the Interservice/Industry Training Simulation and Education Conference, pp. 199-211, 2001.

[3] Choi, I., and Ricci, C. "Foot-Mounted Gesture Detection and its Application in Virtual Environments," Proc. of IEEE International Conference on System, Man and Cybernetics, vol.5, pp. 4248-4253, 1997.

[4] Darken, R., Cockayne, W., Carmein, D. "The Omni-Directional treadmill: a locomotion device for virtual worlds", Proc. of UIST'97, pp. 213-222, 1997.

[5] Ensor, J. R., Carraro, G. U. "Pelton: A VRML-based bicycling simulator", In SIGGRAPH'97 Visual Proc., pp. 198, 1997.

[6] Iwata, H., and Fuji, T. "Virtual perambulator: a novel interface device for locomotion in virtual environment", Proc. of IEEE VRAIS'96, pp. 60-65, 1996.

[7] Bouguila, L., Ishii, M. and Sato, M. "Realizing a new step-in-place locomotion interface for virtual environment with large display system", Eight Eurographics Workshop on Virtual Environments, 2002.

[8] Bouguila, L., Evequoz, F., Courant, M., Hirsbrunner, B. "Walking-Pad: a Step-in-place Locomotion Interface for Virtual Environments", Sixth International Conference on Multimodal Interfaces, Penn State University – USA, October 14th - 15th, 2004.

[9] Noma, H., and Miyasato, T. "Design for locomotion interface in a large-scale virtual environment". ATLAS: ATR Locomotion Interface for Active Self-Motion. DSC-vol.64, Proc. of the ASME Dynamic Systems and control division, pp. 111-118, 1998.

[10] Razzaque, S. Swapp, D. Slater, M. Whitton, M. C. and Steed, A. "Redirected Walking in Place" Eight Eurographics Workshop on Virtual Environments, 2000.

[11] Razzaque, S., Swapp, D., Slater, M., Whitton, M. C., and Steed, A. "Redirected Walking in Place", Eight Eurographics Workshop on Virtual Environments, 2002.

[12] Richard, P., Chamaret, D., Inglese, F. -X., Lucidarme, Ph., Ferrier, J. -L., "Human-Scale Virtual Environment for Product Design: Effect of Sensory Substitution", the International Journal of Virtual Reality, 5(2), pp. 37-44, 2006.

[13] Slater, M., Usoh, M., and Steed, A. "Taking steps: the influence of a walking technique on presence in virtual reality", ACM Trans. On computer-Human Interaction, 2(3), pp. 201-219, 1995.

[14] Templeman, J.N., Denbrook, P. S., Sibert, L. E. "Virtual Locomotion: walking in place through Virtual Environments", Presence Teleoperator and Virtual Environments: Vol. 8, No. 6, pp. 598-617, 1999.

[15] Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., Brooks, F. J. "Walking > Walking-in-place > Flying, in virtual environments". Proceedings of SIGGRAPH'99, pp. 359-364, 1999.

Movement and touch as haptic inputs: Implications for enaction

Barry L Richardson* Mark A Symmons* Dianne B. Wuillemin*
(*)Bionics and Cognitive Science Centre, Monash University, Churchill, Australia 3842
Corresponding e-mail:barry.richardson@arts.monash.edu.au

Abstract

Enactive models focus on sensory-motor links as the building blocks for learning, perception, and even intelligence. This implies a central role for haptics in the linking process, and for kinaesthesis in particular because of its close association with motor activity. Here we describe two experiments in which the contribution of kinaesthetic and cutaneous information are assessed as separate components of haptics. Results indicate that tactile information delivered via static skin can be as or more important than kinaesthetically registered movement.

1. Introduction

One enactivist idea is that intelligence may emerge from sensory-motor links in a manner that can be applied to machines as well as humans. Several theorists (see [1]) have therefore suggested that a machine might learn to be adaptive if it has a way of usefully recording the consequences of what it did (motor output) as a response to a stimulus (sensory input).

The acceptability of this new and exciting approach to understanding intelligence, be it human or machine, depends on a good grasp of the way in which sensory-motor links develop. This in turn implies knowledge of the sensory processes that tell the brain what the motor system has just done, so that the all-important link is possible. What are these sensory processes that facilitate sensory-motor links?

The answer is that all senses offer some feedback of value. Human eyes move independently of the head, and ears move as the head does, and these movements make up some of the kinaesthetic information that is effectively correlated with sensory input as part of the sensory-motor link. But the most direct source of information about the consequences of action is registered by the haptic system. O'Shaughnessy [2] says, for example, "touch is in a certain respect the most important and certainly the most primordial of the senses "(p.658). That the sense of touch as "primordial" is true with respect to both evolution and function. The eye and ear develop embrionically from skin cells. The cutaneous system is better than audition for spatial judgements (but is inferior in temporal resolution), and it is better than vision for temporal resolution (but inferior in relation to spatial judgements). It is the perfect candidate for linking audition and vision within the multimodal context that has to be present as a learning environment for humans or machines.

Also drawing attention to the importance of haptics, Pfiefer and Bongard [1] emphasise its value within the "redundancy" principle which, among others, is critical for effective designs of intelligent systems. Their redundancy principle refers not only to the ability of one sense to compensate for reduced (or absent) fidelity in another, but to the useful *overlap* of information that is possible with multisensory inputs. They say that "...when two modalities yield partially overlapping information…the information extracted from one can be used to…predict the information that can be extracted from the other" (p.114).

Unfortunately, despite a growing recognition of its contribution to perception and cognition, much less is known about haptics than vision and hearing. For instance, although it is generally agreed that haptics is comprised of the kinaesthetic system (which gives information about movement) and touch (cutaneous information), the individual contribution of these senses to haptics, and how their inputs are integrated, remains unclear.

A widely held view has been that since movement is critical for effective perception, kinaesthesis should be more important than touch. This position is central to Gibson's [3] argument that active touch is superior to passive touch. The importance of kinaesthesis was also posited by Magee and Kennedy [4], although they reported better identification of raised-line drawings when subjects' fingers were guided around a drawing, rather being allowed to explore freely. Magee and Kennedy argued that touch merely served to indicate to the subject "you are on the line" or "you are off the line", rather than being an important source of information per se.

Here we report summaries of two experiments that cast doubt on the view that kinaesthetic information is more important than that arising from cutaneous stimulation. For both studies we used a device we call the Tactile Display System, or TDS [5], which records the movements of an explorer's fingertip actively moving around a raised-line drawing, and then reproduces those movements to guide another (passive) subject around the same drawing, matching for position and speed. In addition, the device can be used to move a raised line drawing underneath a stationary fingertip.

Although the TDS has frequently been used to compare how well raised-line drawings can be identified in active and passive conditions [6], in the experiments reported here, subjects were never active, but were either passively guided, or had their finger held still while a stimulus was moved under it.

2. Experiment 1: Systematic degradation of haptic information to compare kinaesthetic and cutaneous contributions

In the first experiment, the amount of haptic information available to subjects was systematically reduced from a maximally informative condition called "full passive-guided", in which the TDS guided the subject's finger over a raised-line drawing, through several conditions to a least informative condition which consisted of a piece of textured paper (with no raised line present) being moved under a stationary fingertip. This last condition was called "shear only" because, with no raised line present, identification of the letter depended on detecting the direction and extent of movement of the textured surface across (under) the stationary fingertip. The pathway this paper followed matched (for position and speed) the pathway followed in the first and all other conditions in between these two extremes. The other conditions of particular interest here were "kinaesthesis only", in which subjects' fingers were guided around the shape of the letter but suspended in the air so that the fingertip did not contact any surface, and "touch only", in which the raised line letter was moved under the subject's stationary fingertip. The latter condition offered information via the raised line and shear forces, but there was no kinaesthetic input.

A pair of stimuli was randomly selected from a set of nine letters (ABGKMQRXZ) for each of eight conditions (only four conditions are discussed here). The 15 blindfolded subjects were told that they would be exploring capital raised-line letters, but were not informed that only nine rather than all 26 letters were used.

Performance was assessed in terms of time taken to identify the letter, and number of stimuli correctly identified. These results are contained in Table 1 for the four conditions of interest.

Table 1. Mean (M) and standard deviation (SD) for latency (Lat), and percentage correct (% corr) as a function of exploratory condition

	Full passive	Kinaesth only	Touch only	Shear only
Mean Lat (s)	13.7	15.1	16.1	19.8
SD Lat (s)	5.9	7.4	6.7	6.8
% corr	85	85	75	35

A repeated measures analysis of variance revealed a significant effect (p<.05) across the four conditions with respect to latency to correct response such that the 15 subjects' mean latency for the full passive condition was 13.7 seconds and the mean latency for the shear only condition was 19.8 seconds. Interestingly, the mean latencies for the kinaesthesis only and touch only conditions (15.1 and 16.1 seconds respectively) did not differ significantly.

A similar pattern of results was obtained for the accuracy data. The mean number of correct identifications was 1.7 (out of a possible 2) in the full passive condition, but 0.7 in the shear only condition. This difference was significant (p<.05). The mean number correct in the kinaesthetic and touch only conditions (1.7 and 1.5 respectively) did not differ significantly.

The finding of particular interest was that kinaesthesis alone did not yield better performance than touch alone, although it could be argued that the trend was in favour of kinaesthesis. To further explore the roles of these components of haptics we conducted a second experiment in which one of the conditions consisted of the simultaneous presentation of stimuli to the fingers of different hands.

3. Experiment 2: Differential attention to kinaesthesis and touch

In this study we used the TDS to hold the finger of a subject's hand steady while a raised line drawing was passed under it, and at the same time, the other finger was guided by the TDS (see Figure 1). The stimuli used were specially chosen so that one finger would experience a 180-degree rotation of the stimulus at the other finger. For example, as shown in Figure 1, the "top tray" on which a raised line drawing of a 9 can be seen, was moved under the fingertip of the left hand at the same time that the fingertip of the right hand was guided over the raised-line drawing of the 6. The other stimulus pairs used were the lower case letters p and d; b and q and the capitals W and M, drawn so that the side strokes of the W were parallel to match the 180-degree rotation of the M.

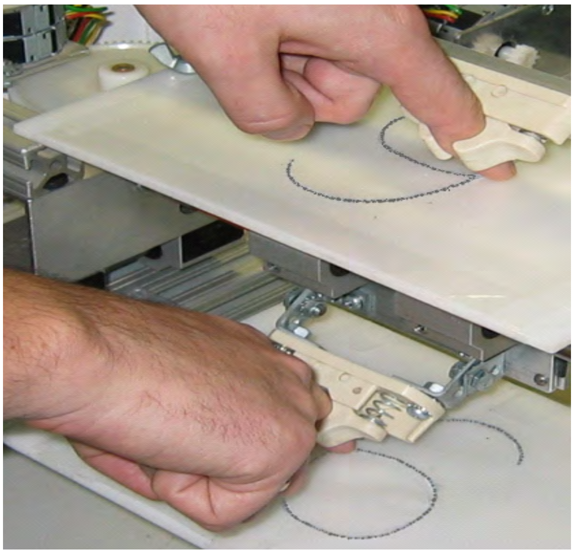


Figure 1. Two-handed simultaneous exploration of 180-degree rotated stimuli using the Tactile Display System. The raised-line 6 on the lower tray is explored by the moving fingertip, while the raised line 9 on the upper tray is moved under a stationary fingertip.

Twenty blindfolded subjects were presented with these four pairs of stimuli, one member of the pair at each fingertip, but were not told of the rotation – instructions gave the impression that the stimuli at both hands were identical (and in fact they were, but rotated versions of one another). We were interested in which stimulus would be reported in the event that subjects did not notice that there were potentially two correct answers. From the report, we would infer the stimulus to which attention was presumably being directed.

When the raised-line 9 was moved under the stationary fingertip of the left hand, and, simultaneously, the fingertip of the right hand was guided over the 6 (as shown in Figure 1) all subjects reported the stimulus at the moving finger – in the specific case shown in Figure 1 the “6” was always reported, suggesting that the “9” at the stationary finger was either ignored, or was “captured” by kinaesthetic and cutaneous information present at the lower, moving hand (see [7] for full results). The same pattern of responding applied to the other pairs of stimuli used -- subjects always reported the stimulus under the guided moving fingertip.

In another condition we removed the tactile stimulus at the moving hand so that as the TDS guided the subject’s finger in the shape of the stimulus (a 6 in the example shown in Fig 1) they did not experience touch or shear information at that finger, only kinaesthesis. The stationary finger still received cutaneous information from the raised line, and shear forces. In this condition, 90% of trials resulted in reports of the stimulus at the *stationary* finger (which was coded by the tactile channel only). The kinaesthetic information, by itself, was apparently not sufficiently attention-attracting to capture touch by itself. This is at variance with the idea that kinaesthesis is the more important component of haptics. The same pattern of results was obtained when subjects were active, that is, when the lower hand (the right hand in Figure 1) was free to explore, instead of being passively guided. These movements caused the upper tray to deliver the 9 raised-line drawing to the stationary fingertip. The only difference between the results in the active and passive conditions was that 100% of subjects reported the stimulus at the stationary fingertip when the moving fingertip received only kinaesthetic information. In the corresponding passive condition, reported previously, 90% of subjects reported the stimulus at the stationery fingertip (where only cutaneous information was present).

Interestingly, none of the subjects in any of the conditions asked which of the two simultaneously presented stimuli should be reported, and none voluntarily reported feeling any more than one stimulus. When prompted to report whether two stimuli had been noticed, only those subjects who had been active, and had felt a raised line at both fingertips, said they had noticed some discrepancy.

4. Conclusions

Taken together, these studies suggest that the contribution of touch (cutaneous information) to haptic percepts has been undervalued in relation to kinaesthetic input. Magee and Kennedy's [4] claim, that the role of touch was primarily to inform kinaesthesis by indicating presence or absence of the line under the fingertip, is challenged by our findings. So too is Noe's [8] claim that touch can tell us virtually nothing without movement, "It seems plausible", he argues "that feeling alone is not sufficient to enable you to learn about or discover the properties of objects or layouts around you", and also asks "How could you perceive [an] object *as rectangular* without moving it across your body surfaces, or without moving your body surfaces across it” (p.15). According to our results, the answer to this question could be "by having the stimulus moved across your body surfaces by another agent”.

The movement of the stimulus under the stationary fingertip in our studies was sufficient for identification of the stimulus, even when no line could be felt (as was the case in the first experiment described above) so that judgements depended on shear forces alone. For most enactivist models, kinaesthetically registered movement (e.g., that arising from activity) is crucial for the development of sensory-motor links that may be the foundation of perception. These models do not appear to include movement across the skin induced by an outside agent as useful from a sensory-motor standpoint. Accoring to Noe the enactivist view which "insists that mere feeling is not sufficient for perceptual experience" (p.16). Having another agent move something across your skin would appear to qualify as "mere feeling", and should not, therefore, have resulted in a perceptual experience for our subjects.

Of course, it could be argued that previous tactile experience of shapes has involved self-induced movements that have allowed strong percepts to be built by the sensory-motor links so formed. When only tactile information is available, that learning comes into play and the percept is triggered as if the self-induced movement was still present as an input. Recent evidence suggests that not only is this possible, but that multimodal integration allows such association at a very early stage of processing, and without the need for sensory-motor links to be slowly learned. Research on bimodal cells in monkeys has shown that there are touch-responsive cells in the parietal lobe that also respond to visual information indicating that something is *about* to touch the part of the body served by those cells [9]. Specifically, there is a three-dimensional "peripersonal space" adjacent to tactile receptive fields and if an object can be seen as it enters this region, the associated cortical cells respond as if touched. Importantly, the visual area (receptive field) next to the tactile receptive field, moves with the body part, and not the eye. Cells have also been found that respond when an action is performed, when a sound related to that action is heard, or when that action is merely seen [10].

These so-called "mirror" neurons, and the bimodal neurons described earlier, are illustrations of multisensory convergence at an early stage of processing, and they offer ample support for enactivist models without there having to be dependence on motor activity for perceptual development. This suggests that kinaesthesis need not be seen within enactivist models as the most important information-carrying component of haptics. Instead, sensory-motor links may benefit just as much from inputs through static skin, which our results suggest to be more useful than previously thought.

References

[1] Pfiefer, R., and Bongard, J. How the Body Shapes the Way We Think, MIT Press, Cambridge, MA, 2007.

[2] O'Shaughnessy, B. Consciousness and the World, Oxford University Press, New York, 2000.

[3] Gibson, J.J. The Senses Considered as Perceptual Systems, Houghton Mifflin, Boston, 1966.

[4] Magee, L.E., and Kennedy, J..M. "Exploring Pictures Tactually", Nature, 283, p. 287, 1980.

[5] Richardson, B.L., Symmons, M.A., and Accardi, R. "The TDS: A New Device for Comparing Active and Passive-guided Touch," IEEE Trans. Rehab. Eng., 8, pp. 414-417, 2000.

[6] Symmons, M.A., Richardson, B.L., and Wuillemin, D.B. "Active vs Passive Touch: The State of Play and the Future" IEEE Proc. World Haptics, Tsukuba, Japan, pp567-568, 2007a.

[7] Symmons, M.A., Richardson, B.L., and Wuillemin, D.B. "Intrasensory Attention: Kinaesthetic versus Cutaneous Inputs", Perception, 36, pp. 880-887, 2007b.

[8] Noe, A. "Action in Perception", MIT Press, Cambridge, MA, 2004.

[9] Graziano, M.S.A., and Gross, C.G. The representation of extrapersonal space: A possible role for bimodal visual-tactile neurons, In M.S.Gazzaniga (Ed.), The Cognitive Neurosciences, pp 1021-1034, MIT Press, Cambridge, MA, 1995.

[10] Fu, K-M.G., Johnston, T.A., Shah, A.S., Arnold, L., Smiley, J., Hackett, T.A., Garraghty, P.E., and Schroeder, C.E. "Auditory Cortical Neurons Respond to Somatosensory Stimulation" Journal of Neuroscience, 23 (20), pp.7510-7515, 2003.

Tongue Liminary Threshold Identification to Electrotactile Stimulation

Robineau Fabien* Vuillerme Nicolas* Jean-Pierre Orliaguet*† Yohan Payan*
(*)TIMC IMAG Laboratory, CNRS-UMR 5525, France
(†)LPNC Laboratory, CNRS-UMR 5105, France*
E-mail: Fabien.Robineau@imag.fr

Abstract

Many applications use electrostimulation of the human skin to provide tactile sensation. The effect of electrotactile stimulations were studied on a 6x6 matrix of tactile electrodes placed on the anterior part of the tongue. The liminary threshold with continuous or discontinuous waveform and patterns with 2 or 4 electrodes was investigated.

The result suggest that for energy saving and to improve the yield, it would probably be better to use discontinuous stimulation with two electrode patterns.

1. Introduction

Electrotactile stimulation can provide sensory information not available to damaged classical sensory system (vision, hearing, vestibular or somatosensory systems). Many electrotactile human machines interface (HMI) have been developed and applied to various surface areas [2]. The first development of electrotactile visual substitution systems (TVSS) was designed to provide distal spatial information to blind people via a 20x20 electrodes matrix [5]. Later Bach-y-Rita and his collaborators converged towards the electro-stimulation of the tongue surface, called "Tongue Display Unit" (TDU), which provides a practical HMI [4]. The human tongue is very sensitive, highly mobile and discriminative (spatial threshold < 2 mm). Moreover, the environment of the mouth offers a protected volume.

As concerns the tongue sensitivity, most authors agree that the tactile sensitivity threshold of the anterior part of the tongue, and particularly the tip, is lower than the other part [10, 9]. Moreover, it was showed that the mechano-receptive innervation is denser on the tip than in the other regions. However, few researches have been investigated about the asymmetry of tongue sensitivity with respect to the median sulcus (right and left tongue sides) [13].

TIMC-IMAG laboratory recently developed medical electrotactile human machine interfaces using the TDU of Bach-y-Rita with a 6x6 matrix. These devices, which feasibility was evidenced, are used for disable persons assistance and for computer-aided surgery. Either the TDU provides electrotactile signal to supply punctual information, or the functioning is almost continuous to give orientation information. An application to prevent pressure ulcer formation in paraplegics, developed as an “alarm”, provides via the TDU, tactile information about excess of pressure at the skin/seat interface [11]. In addition, a device to improve human ankle joint position sense was investigated [16]. On the other hand, electrotactile information comparable to continuous information is investigating. In this way, a biofeedback system aims at improving human balance control for people with visual or hearing impairments [15]. Also, is studied a system to provide to a surgeon, via the TDU device, orientation information to accuracy guide a needle until a target inside a body [14].

In these studies, a minimum of four electrode patterns were used because the subjects were able to discriminate without problem the stimuli. In parallel, a more ergonomic wireless 6x6 TDU device was recently developed in our laboratory. It is inserted in a dental retainer including microelectronics, antenna and power supply [14]. To increase the comfort of such an embedded device, the size and the energy consumption have to be even more reduced. The intensity on the matrix surface has to be as low as possible, but of course sufficient to be perceived by the subject (i.e. higher than the “liminary threshold” of the human tongue). Another option to decrease such intensity would be to reduce the number of activated electrodes. However, such a reduction should not decrease the discrimination of the electrotactile patterns.

The present study aims at evaluating the liminary thresholds of the tongue with continuous or discontinuous electrotactile stimulations among subjects, and comparing such thresholds with 2 or 4 electrode patterns.

2. Methods

2.1 Subjects

Two groups of ten subjects (age: 28,3 ± 3,9 years; body weight: 68,8 ± 10,2 kg; height: 174,6 ± 9,9 cm) voluntarily participated in this experiment. None of the subjects presented any history of sensory/motor or cognitive problem.

2.2 Apparatus

Electrotactile stimuli were delivered to the dorsum of the tongue via a ribbon TDU derived from the one developed by Bach-y-Rita [2, 4]. This electrotactile device consisted of an array of 36 tactile electrodes (6×6 matrix, radius: 0.7mm each), embedded in a 1,5×1,5 cm plastic strip (Fig. 1, left panel). The tip of the TDU was inserted in the oral cavity and held lightly between the lips, maintaining the array in close and permanent contact with the surface of the tongue (Fig. 1, right panel). A flexible cable, made of a thin (100 µm) strip of polyester material, connected the matrix to an external electronic device. This device delivered the electrical signals that activated the tactile receptors on the anterior superior part of the tongue. As Kaczmarek suggested [8], the frequency of the DC stimulating pulses was fixed to 50Hz for all trials and subjects. Because of the conductive properties of the saliva and the epidermis thickness, the TDU only required a 5-15V input voltage and a 0.4-4.0mA current. Each gold-plated circular electrode (Ø = 1.4mm, inter-centre distance = 2.3mm) received monophasic pulses.

Fig. 1 On the left panel, the Tongue Display Unit. It constitutes of 2D electrode array (1,5 cm × 1,5 cm) including 36 gold-plated contacts each with 1.4 mm diameter, arranged in a 6×6matrix. On the right panel, the TDU placed on the tongue.

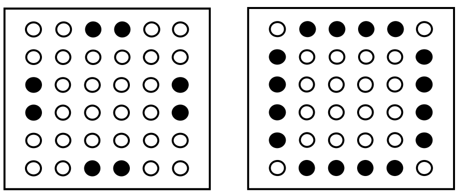


Fig. 2 Electrode Schemes of the four patterns couples tested on the anterior part of the tongue. Left panel: patterns with 2 electrodes; right panel: patterns with 4 electrodes.

Considering that each electrode could be independently activated, four tactile patterns couples were designed. As illustrated in Fig. 2, these patterns consist in the activation of either 2 or 4 electrodes located on the front, the left, the back and the right side of the matrix.

2.3 Task and procedure

Before the experiment, a set of instructions describing the different kind of patterns used in the study and the required task was orally presented to each participant. The TDU matrix was placed inside the mouth and in contact with the anterior superior surface of the tongue. The participants were asked to keep the TDU in the same position on her/his tongue as well as possible during the course of the experiment.

Two groups carried out the experiment. The first group performed the test with continuous waveform (CT condition) and the second group with discontinuous waveform (DCT condition). For each group, the patterns with two electrodes were first tested, followed by the patterns with four electrodes.

A psychophysical technique referred to as the "method of limits" [7] was employed in determining each pattern's liminary thresholds for each subject. An ascending threshold was established by beginning with a very low, previously tested stimulation threshold, gradually increasing by steps of 150 mV until the subject perceived a stimulation. A descending threshold was established by beginning with a stimulation clearly detectable (previously tested) and gradually decreasing by steps of 150 mV until the subject could no longer perceive any stimulation. A dedicated software administered the test of the liminary threshold detection. A total of 5 ascending and 5 descending series were alternatively run to each of the eight patterns for each subject. During the continuous stimulation test, the stimuli intensity was maintained 3s then increased or decreased as a function of the phase (ascending or descending). In the case of the discontinuous stimulation, the stimuli were sent during 1 s every 3 s in order to avoid a sensory adaptation and evoke a good recovery.

The subject was required to signal when she/he was absolutely certain of the appearance (ascending test) or the disappearance (descending test) of the lingual stimulation, by pressing a user-defined key on the laptop placed in front of her/him. To indicate the direction of the test, an up or down arrow was shown on the computer screen. A red point was also shown to specify when there was stimulation. Each testing session lasted about 2 hours.

2.4. Data analysis

The experiment yielded a total of 100 thresholds per pattern tested. The identified threshold for each pattern and for each subject was defined as the average of the median value for the 5 ascending responses and the median value for the 5 descending responses. The threshold values were calculated as the value at 50% of the 5 ascending or descending responses, with a linear interpolation supposing that the probability variation was linear between two values which surrounded 50% .

Data were computed and treated with Statistica statistics software. They were then submitted to four separated 2 Groups (Continuous vs Discontinuous) x 2 Electrode Patterns (2 electrodes vs 4 electrodes) analyses of variances (ANOVAs). The level of significance was set at 0.05.

3. Results

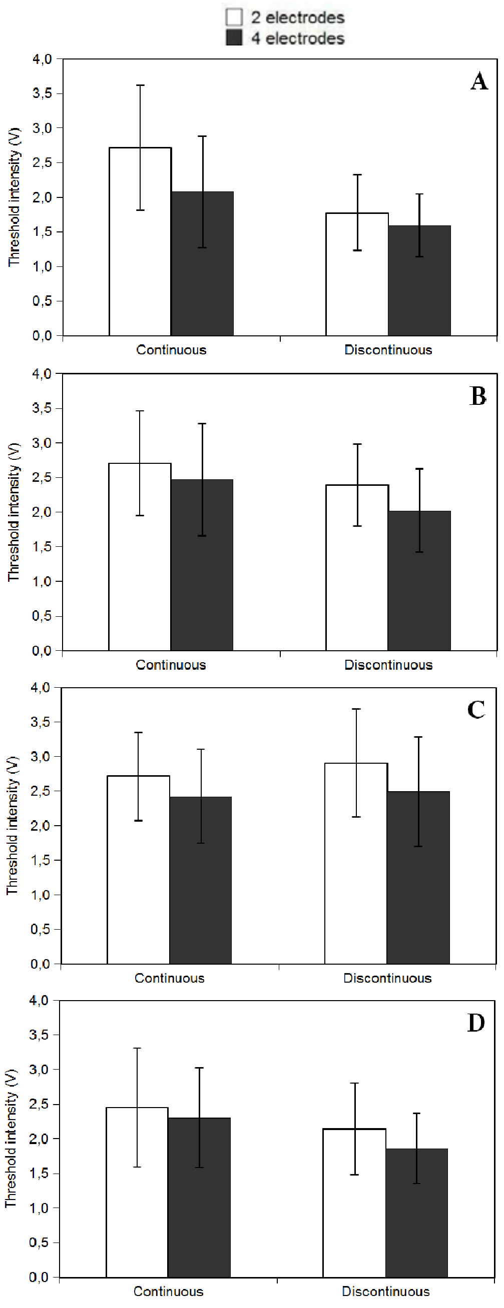


Fig. 3 Mean intensity liminary threshold for CT and DCT condition as a function of electrode number, 2 or 4, for A) the front, B) the left, C) the back, D) the right positions.

The mean intensity liminary threshold (± standard deviations) were obtained for CT and DCT conditions as a function of number of activated electrodes (2 or 4), for the four front (A), left (B), back (C) and right (D) positions.

The ANOVA shows a significant interaction between the groups and electrode numbers conditions for the front position (Fig. 3A), (F (1;18)=6.22, $p < 0.03$). Planned comparison analyses evidence a significant difference between two and four electrodes for continuous stimulation (F (1;18)=24.48, $p < 0.001$), whereas no significant difference was observed for the discontinuous stimulation (F (1;18)=2.01, $p > 0.17$).

The ANOVA also shows main effects for CT and DCT group, for the front position (F (1;18)=15.01, $p < 0.01$, figure 3A), and electrode patterns condition, for the left, back and right positions (respectively (1;18)=13.33, $p < 0{,}01$; F (1;18)=18.68, $p < 0.001$ and F (1;18)=16.13, $p < 0.001$). Liminary threshold is higher for the continuous than for the discontinuous condition and the patterns with 2 electrodes require an intensity threshold higher than the patterns with 4 electrodes.

4. Discussion

Tactile vibratory sensations are transmitted on the tongue surface via electrotactile stimulation. The electric current passing through the skin stimulates cutaneous afferent fibers at the location of the tactile electrodes. Our laboratory addressed various medical applications using the electrotactile feedback to supplement the damaged sensory information or provide accurate orientation information not easily accessible.

The front part of the tongue shows a significant difference for the continuous vs discontinuous conditions. As previously reported in the literature [9] [10] [12] [13], the sensitivity of the tongue tip is higher than all the others parts. Interestingly, our result further demonstrated that on the tip tongue, the liminary threshold is higher using continuous stimulations than discontinuous. This high tip tongue sensitivity of the tongue anterior part can explain that the habitutation effect is faster with the continuous than with discontinuous stimulations. Indeed, after habituation, the induced sensation becomes less clear and less discriminable. Therefore, the performances of the subjects are reduced. An option to limit this adaptation could be to find the best stimulation frequency to improve the subjects performance (alarm or continuous guidance) and to optimize the tongue tactile neuronal adaptation/recovery in order to allow a regular use.

In this experiment, all the electrode number main effects were significant. For each position (front, left, back and right), the electrotactile liminary threshold was higher for two electrodes than for four electrodes. This result is quite natural and consistent with the literature [1]. However, it is interesting to compare the reduced intensity saved with respect to the electrode number activated to optimize the energy consumption. Although the two and four electrode patterns intensity liminary thresholds are very closed (only a 0,9 ratio between both), there are twice more activated electrodes between both patterns. Consequently, patterns with two electrodes should be preferred to patterns with four electrodes, assuming that the subjects are still able to efficiently perceive such patterns.

Moreover, the observation of a significant interaction between both conditions (CT/DCT vs Electrode number) on the front part shows that the liminary threshold difference between two and four electrode patterns observed for the continuous stimulation is not present under discontinuous stimulation. These results suggest that, for energy saving and to improve the yield, it would probably be better to use discontinuous stimulation with two electrode patterns.

Obviously, complementary studies should be performed to investigate more quantitatively lingual electrotactile stimulation. Psychophysical experiments have shown that it possible to provide electrotactile sensation on the tongue with a very light current. Also, further experiments could allow optimising the comfortability threshold of the wireless TDU for a daily use.

Acknowledgement

This research was supported by the Gravit ("Grenoble Alpes Valorisation Innovation Technologie") consortium and the “Multimodal Interaction program” of Grenoble University.

References

[1] Ajdukovic, D., "The relationship between electrode area and sensory qualities in lelectrical tongue stimulation", Acta Otolaryngol, Stockholm, 1984.

[2] Bach-y-Rita, P., Tyler, M.E., and Kaczmarek, K.A., “Seeing with the brain,” Int. J. Hum-Comput. Int., Vol. 15, pp. 285-295, 2003.

[3] Bach-y-Rita, P. and Kaczmarek, K., “Tongue Placed Tactile Output Device,” U.S. Patent 6,430,450 B1, 2002.

[4] Bach-y-Rita, P., Kaczmarek, K.A., Tyler, M.E., and Garcia-Lara, J., “Form perception with a 49-point electrotactile stimulus array on the tongue,” J. Rehabil. Res. Dev., Vol. 35, pp. 427-430, 1998.

[5] Bach-y-Rita, P., Collins, C.C., Saunders, and al., "Vision substitution by tactile image projection", Nature, Vol. 221, pp. 963-964, 1969.

[6] Collins, C. and Bach-y-Rita, P., “Transmission of pictorial information through the skin,” Advances in Biology and Med. Phys., Vol. 14, pp. 285-315, 1973.

[7] Gescheider, G.A., "Psychophysics: The Fundamentals", Lawrence Erlbaum Associates, Mahwah, NJ, 1997.

[8] Kaczmarek, K.A., “Electrotactile adaptation on the abdomen: Preliminary results,” IEEE Trans. Rehab. Eng., Vol.8, pp. 499-505, 2000.

[9] Maeyaman T., Plattign K.H., "Minimal two-point discrimination in human tongue and palate", Am J Otolaryngol, Vol. 10, pp. 342-4, 1989.

[10] Marlow, C.D., Winkemann, R. K., Gibilsco, J.A. General sensory innervation of the human tongue. Anat. Rec., Vol. 152:, pp. 503-512, 1965.

[11] Moreau-Gaudry, A, Prince, A., Demongeo,t J. & Payan Y., "A New Health Strategy to Prevent Pressure Ulcer Formation in Paraplegics using Computer and Sensory Substitution via the Tongue", Studies in Health Technology and Informatics, Vol. 124, pp. 926-931, 2006.

[12] Pleasonton, A.K., "Sensitivity of the tongue to electrical stimulation", J Speech Hear Res., Vol. 13(3), pp. 635-44, 1970.

[13] Ringel, R. L. and Ewanowski, S.J., "Oral perception. I. Two-point discrimination", J. Speech Hear. Res., Vol. 8, pp. 389-398, 1965.

[14] Robineau, F., Boy, F., Orliaguet, J-P., Demongeot, J. & Payan, Y. "Guiding the surgical gesture using an electro-tactile stimulus array on the tongue: A feasibility study", IEEE Transactions on Biomedical Engineering, Vol. 54(4), pp. 711-717, 2007.

[15] Vuillerme, N, Chenu, O, Demongeot, J. & Payan, Y. "Controlling posture using a plantar pressure-based, tongue-placed tactile biofeedback system", Experimental Brain Research, Vol. 179, pp. 409-414, 2007.

[16] Vuillerme N., Chenu O., Demongeot J. & Payan Y., "Improving human ankle joint position sense using an artificial tongue-placed tactile biofeedback", Neuroscience Letters, Vol. 405, pp. 19-23, 2006.

Performer as perceiver: perceiver as performer

Matthew Rodger Johann Issartel Sile O'Modhrain
SARC, Queen's University, Belfast
E-mail: mrodger01@qub.ac.uk

Abstract

The proposition in the title of this paper is intended to draw a link between psychological processes involved in aesthetic gestural performance (e.g. music, dance) for both performers and perceivers. In the performance scenario, the player/dancer/etc., perceptually guides their actions, and acquires the skill for a performance through their previous perceptions. On the other side, the perceiver watching, listening to and experiencing another's motor performance, simulates the actions of the performance within the range of their own motor capabilities. These phenomena are possible due to common mechanisms of action and perception, and in tandem provide the basis for the rich experience of gestural performance.

This paper reviews evidence for these claims, using examples from the domains of music and dance performance. Questions that arise from these propositions are addressed and suggested empirical explorations of these ideas are given. Further problems in incorporating these theories about gestural performance experience within Enaction are highlighted for future discussion.

1. Introduction

This review puts forward the concept that gestural performance experience arises through common mechanisms of perception and action. Initially, it will be beneficial to clarify use of the term 'gesture'. Firstly, research and theory about 'action' are here applied to gesture, implying equivalence between the two terms. As intention is a necessary component of action [20], in distinguishing action from mere movement, the implied equivalence between gesture and action here may raise issues about the role of intention in performance gestures. To be clear, the main intention is considered to be the performance intention as a whole, to which other actions are considered to involve more basic intentions (see [10]). Regarding ancillary gestures in music performance, for example, the main intention of the performer may not necessarily be ancillary gestures per se, but rather the intention to communicate a holistic musical 'expressive unit' [26], for which the ancillary and effective gestures are more basic, subsumed within the overall performance intention. Secondly, regarding musical performance, it is considered that perception of auditory events will induce simulation of gestural action thus allowing for non-visual communication of action. This is emphasised here as discussion of movement in dance and ancillary gestures in music performance may bias understanding towards visual communication in performance and this is insufficient for understanding musical performance as a whole.

.

2. Performer as perceiver

The 'performer as perceiver' can refer to two features of gestural performance. Firstly, the performer is engaged in perceptual guidance of their movements during the actions that constitute performance [13], and can to an extent modify their actions on-line on the basis of perceptual feedback and attention [21]. Thus, the performer is an active perceiver of their own bodily movement and bodily interaction with instrument, through the gestural channel [2], or interaction with conspecifics [8]. Secondly, the extent to which performance actions are not available to on-line guidance and manipulation delimits the extent to which memorised or pre-programmed movement is necessary for performance. It is here that the acquisition of performance content and skill is effective. Present skill level and characteristics will be formed through prior perception and feedback, both of one's own performances, through the development of sensorimotor loops, as well as from observational effects, either from instructive demonstration [9] or in audience of related performances. It thus seems that both during performance and during its acquisition, the music or dance is actively shaped by current and previous perception.

In musical performance specifically, empirical support for the performer as perceiver comes from a number of sources. Evidence shows that effective musical movements in performance are perceptually guided, for example in the vocal closure of pitch gaps

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

[13]. However, the connection between auditory perception and effective sound-producing movement seems to be more salient in skill acquisition than in online guidance. This is illustrated in studies where altered auditory feedback had a stronger effect on memorisation of a musical piece [5] and on development of expressive variation [15], than on online performance errors.

At a more holistic level of musical performance, the presence of ancillary gestures as communicative cues [26] indicates that performers are perceptually sensitive to effective means of communication in performance, as those gestures that are felt to fit the music become incorporated into their physical performance. This is neatly summarised by Thompson et al. [24]: "visual aspects of performance signal that performers are not merely producers of sound but are themselves listeners, highlighting the musical activity as a shared experience between performers and listeners"–p178. To explore the nature of this developmental aspect of performance, research is in progress to explore the ancillary gestures made by musician as a function of skill acquisition. By comparing the presence and characteristics of ancillary gestures in performers of different skill levels, and over the acquisition of a new piece of music, the effects of self-recursive perception on these performance gestures may be illuminated. It is hypothesised that as skill level increases, stability and consistency of gestural movements will increase, and that the relationship between gestures and musical parameters will become stronger, as gestures come to 'fit' the musical intention.

This point leads to the question of whether it is possible to consider ancillary gestures in music as dance performance, where there is an explicit focus on the 'fit' between movement and performance intention. In the case of a positive response, the question arises of what could be the weight of such movement in the musical performance of the performer? In the same way but from the perceiver point of view, what could be the consequences on the perception of the musical performance?

3. Perceiver as performer

As with the first proposition of the title, the second can refer to two aspects of the perception of gestural performance. The first concerns the contention that embodied motor simulation is central to perception of the actions of others [11], and by extension, perception gestural performance will entail embodied simulation of those gestures. To the extent that embodied simulation of gesture obtains, the perceiver is a covertly simulated performer in resonance with the actual unfolding performance. A corollary of this idea lies in the boundaries of its limitations. A perceiver does not actually realise the actions in simulation, a fact which is necessary to maintain the self/not-self distinction [11], that is, 'I am resonating with the performer, but I am not *myself* the performer'. This trivial fact becomes less trivial when looking at studies of action identities in, for example, piano performance. The ability of pianists to make more accurate judgements about their own performance than others [16] may likely lie in the greater resonance potential between their own motor system and that employed in the stimulus performance, i.e., their own recording. This observation leads to the further possibility that the degree of communication and engagement between performer and perceiver in a natural setting may be partially determined by the match between their respective motor capabilities.

Secondly, due to differences in motor mechanisms between skilled and non-skilled performers within a given domain, it follows from this account that there will be differences between skilled and non-skilled perception of a given performance. That is, the extent to which a perceiver is capable of simulating performance actions, will be a function of their own action capabilities, and in turn will shape their own perception of a performance, with skilled perceivers being closer matched to the performers. For example, the observation that trained musicians are greater able to articulate in verbal report their conscious experience of music as evidence of a richer conceptual content of their experience [1], coupled to the claim that concepts have a basis in the motor system [7], indicates the dependency of musical perception on motor capabilities.

An empirical imperative that follows from this possibility is to examine the extent to which skilled performers differ from non-skilled in their perceptual judgements of different performances. Some evidence to answer this question has been gathered in studies of dancers. One study in progress, conducted by one of the authors, has analysed real-time judgements of dancers by non-dancers Preliminary results revealed that although global choreography judgments may be different, participants followed the same relative judging pattern. This suggests that, as in tasks where the subject is performing the behavior, so too in judgements of motor behaviors we are naturally attracted by certain movements. Another study from the neuroscience literature compared fMRI activity of dancers watching performances of their own dance genre with performances from another genre [3]. The results showed greater activity in the premotor cortex when witnessing a dance genre that they have acquired greater motor capabilities in. This indicates that familiarity with motor skills of a specific genre affects processing of dance performance. A further study of perception of gymnastics found that skill level of perceiver influences judgements, and hence perception, of both aesthetics and skill in the gymnast performance [19]. In order to extend this last finding to the domain of music, a study to be conducted will compare the ratings of musicians of different skill levels in response to musical performances by players of different skill levels. This is intended to uncover whether musicians are more consistent in their perceptions of performances by players of a similar skill level, and hence motor capability, as themselves. As a result, this research will explore the effects that motor ability compatibility has on perception, as an indicator of the resonance potential of simulated gestural performance.

4. Exploring the flexibility within common performance-perception experience

The claim that a person's perception of gestural performance can depend on their motor capabilities could be taken to mean that only skilled performers can be engaged by a performance, and only if it is in their relevant skill domain. Schutz-Bosbach & Prinz [18] imply this when they suggest that, "[i]f you have never played tennis in your life, you will probably never think of buying a ticket to watch a game at Wimbledon." – p349. This is not the claim we wish to make, as gestural performance has a universal attraction to people with different motor systems, as shown by the universality of music and dance [14]. Rather, the differences in perception between skilled and non-skilled performers that result from differences in respective motor capabilities are better considered as smaller individual variations within a larger common motion-perception spectrum, which is grounded in our physical and physiological similarities.

A motion-perception spectrum runs from the perception of non-biological movements at one end to the perception of self-actions at the other. Progression up the spectrum might correspond to the degree of activation of the common mechanisms of perception and action that develops strength through familiarity with one's own actions in and of the environment. At the lower end, it is observed that perception of non-biological events is constrained by the motor system's sensitivity to biological motion, in preservation of the 2/3 power law [25], or perception of biologically plausible movements in non-biological stimuli [23]. Additionally, it has been proposed that perception and prediction of inimitable events in nature involves attenuated simulation in the motor mirror system of possible actions that share sufficient motional features of the events [17]. The other end of the spectrum is characterised by research into self-action identities that reveal a greater perceptual sensitivity for one's own movements as they have a greater correspondence to one's own motor system capabilities [12]. Where our individual perceptual experience of an action will lie on the motion-perception spectrum will partially be a function of our familiarity, which relates to our motor resonance, with the motion or action perceived, but will have greater dependency on our physiology, which has greater similarity than differences between people. A further observation of relevance is that in the mirror system, which is implicated in perception of other's actions [11], ~30% of mirror neurons are responsive to perception and execution of actions that identical, whereas ~70% are responsive to any similar or related action [7]. In terms of the motion-perception spectrum, this suggests that motor-capability-dependent differences between peoples' perceptions of action are more likely to result from differences in the smaller percentage of neurons, as these are involved in actions that can be identically simulated, whereas the greater number can be approximately simulated, facilitating perception of skilled gesture by non-skilled perceivers.

The common motion-perception spectrum can be related to gestural performance by considering the following example. Schubotz [17] describes listening to a performance by the pianist Glenn Gould as events supposedly inimitable to the majority of perceivers and presumably mediated by the perceiver's ability as a pianist, familiarity with the piece, etc. With this example of virtuoso music, while it is not possible for a non-musician to reproduce the effective gestures of the great player, it is still possible for them to reproduce the kinematics of the holistic (musical) gestures that envelope individual effective gestures [6]. Following Grush's emulation theory, Clark [4] comments that "[w]hen we hear the beat, we are implicitly aware of our capacity, should we wish, to tap our fingers in time with the pulses, to anticipate the pulses, to swing a conductor's baton in time with the pulses, and so on"-p35. Extending this to the case of the virtuoso performance, there are congruent gestures that are possible to make, which correspond to gestures in the virtuoso performance at a level higher than individual notes say, but are rather more like the gestures of a conductor, or ancillary gestures of a performer. Indeed, this may provide the basis to the communicative powers of ancillary gestures in music performance, as they can facilitate attention towards the imitable movements in the musical sound. As Bharucha et al. [1] observe, there are different perceptual experiences of motion available in music and while non-skilled perceivers may not be able to simulate the individual effective motor skills, they are able to access the higher-order motion that may be the vehicle for expressive communication (see also, [22]).

5. Conclusion and further questions

The exposition in this paper of performer as perceiver and perceiver as performer claims that common mechanisms in perception and action are the vehicles for experience of gestural performance. Performance actions are guided and shaped in time by perception of oneself and others. Conversely, perception of performance gestures involves the covert simulation of those same gestures, to the extent that this is possible given the perceiver's motor capabilities. Studies of skill acquisition in gestural performance may provide a means to explore the effects that motor capability have on the perception of performance. However, the differences in the motor capabilities of people are smaller than the similarities. It is this point that endows music and dance with their universality. Hence, in exploring the effects of skill and motor ability on perception of gestural performance, we are exploring variations that exist on a wider common perceptual spectrum.

A number of questions arise in integrating the theory outlined above with broader theories of enaction. Firstly, how to characterise the neural mechanisms of perception and action implicated in gestural performance without positing mental representations of action as other authors do [11]. Another issue is the extent to which motor simulation and acquisition of motor capabilities that form the basis of perception of gestural performance in this story can be considered as enactive knowledge. These questions represent the problems for relating the theory in this paper to wider theories of enaction and while it is beyond the scope of this work to attempt to answer them now, it is believed that the theoretical and empirical issues discussed above will be informative in undertaking this task.

References

[1] Bharucha, J. J., Curtis, M. and Paroo, K. "Varieties of Musical Experience", Cognition, 131-172, 2006.

[2] Cadoz, C. and Wanderley, M. M. "Gesture-Music" In Wanderley, M. M. and Battier, M. Trends in Gestural Control of Music, IRCAM, Paris, 2000.

[3] Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E. and Haggard, P., "Action Observation and Acquired Motor Skills: An fMRI Study with Expert Dancers", Cerebral Cortex, 1243-1249, 2005.

[4] Clark, A. "Visual Experience and Motor Action: Are the Bonds Too Tight?", Philosophical Review, 1-41, 2001.

[5] Finney, S. A. and Palmer, C. "Auditory Feedback and Memory for Music Performance: Sound Evidence for an Encoding Effect", Memory and Cognition, 51-64, 2003.

[6] Friberg, A. and Sundberg, J. "Does Music Performance Allude to Locomotion?", Journal of the Acoustical Society of America, 1469-1484, 1999.

[7] Gallese, V. and Lakoff, G. "The Brain' s Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge", Cognitive Neuropsychology, 455-479, 2005.

[8] Hieronimus, G. Y a-t-il une rupture entre musique du passé et musique d'aujourd'hui? In E. Sprogis, Musique et Dance Contempo…Reines: 1987-1997. Poitiers: Conservatoire National de Région de Poitiers, 1998.

[9] Horn, R. R. and Williams, A. M. "Observational Motor Learning: Is It Time We Took Another Look?" In Williams, A. M. and Hodges, N. J. Skill Acquisition in Sport: Research, Theory and Practice, Routledge, London, 2004.

[10] Hurley, S. L. Consciousness in Action, Harvard University Press, Cambridge, London, 1998.

[11] Jeannerod, M. Motor Cognition, Oxford University Press, Oxford, 2006.

[12] Knoblich, G. and Flach, R. "Predicting the Effects of Action: Interactions of Perception and Action", Psychological Science, 467-72, 2001.

[13] Lee, D. N. "How Movement is Guided". Submitted.

[14] Mithen, S. The Singing Neanderthals, Harvard University Press, Harvard, 2005.

[15] Repp, B. H. "Effects of Auditory Feedback Deprivation on Expressive Piano Performance", Music Perception, 409-438, 1999.

[16] Repp, B. H. and Knoblich, G. "Perceiving Action Identities: How Pianists Recognize Their Own Performance", Psychological Science, 604-609, 2004.

[17] Schubotz, R. I. "Prediction of External Events with Our Motor System: Towards a New Framework", Trends in Cognitive Sciences, 211-218, 2007.

[18] Schutz-Bosbach, S. and Prinz, W. "Perceptual Resonance: Action-induced Modulation of Perception", Trends In Cognitive Sciences, 349-355, 2007.

[19] Scully, D.M., "Visual Perception of Technical Execution and Aesthetic Quality in Biological Motion", Human Movement Science, 185-206, 1986.

[20] Searle, J. R. Intentionality, Cambridge University Press, Cambridge, 1983.

[21] Shaffer, L. H. "Performances of Chopin, Bach and Bartok: studies in motor programming", Cognitive Psychology, pp. 326-376, Elsevier, Amsterdam, 1981.

[22] Sheets-Johnstone, M. "Emotion and Movement", Journal of Consciousness Studies, 259-277, 1999.

[23] Shiffrar, M. and Freyd, J. J. "Apparent Motion of the Human Body", Psychological Science, 257-264, 1990.

[24] Thompson, W. F. and Graham, P. and Russo, F. R. "Seeing Music Performance: Visual Influences on Perception and Experience", Semiotica, 177–201, 2005.

[25] Viviani, P. and Stucchi, N. "Biological Movements Look Uniform. Evidence of Motor-Perceptual Interactions", Journal of Experimental Psychology: Human Perception and Performance, 603-623, 1992.

[26] Wanderley, M. M. and Vines, B. "Origins and Functions of Clarinettists' Ancillary Gestures". In Gritten, A. and King, E. Music and Gesture, Ashgate, Aldershot, 2006.

Inventa sunt specula ut homo ipse se nosset

Francis Rousseaux‡
Université de Reims Champagne Ardenne, France
E-mail: francis.rousseaux@univ-reims.fr

Abstract

When we specify enactive systems leading to an immersion into virtual worlds, we find it natural to give priority to solutions of cognitive continuity, supposed to make the crossing of real/virtual borders easier[1]. Thanks to clever devices carefully elaborated, the users can get the illusion of a perfect ambivalence of interactive relations they are developing in real/virtual environments.

We are thus in favour of user-friendly systems, whose appropriation should thus be obvious to all.

This hypothesis is so widespread and so rarely explained that it remains widely un-thought. However, examples of person/device couplings are numerous, and even though they are restrictive, they remain nonetheless attractive or fascinating: violin or golf players can testify to this.

We suggest discussing the surprising appropriation we spontaneously have of such a common and stupefying instrument as a mirror. Indeed, mirrors play a very important role in our perception of the numerous movements of our body, for instance when we are shaving or putting on make-up.

1. The surprising example of our mirrors

Let us now consider these familiar instruments we use in our everyday lives, whoever and whatever we are, either a dancer, a hairdresser, a car driver or simply a reader of Enactive'07.

Nothing will ever prevent a young woman from using her pocket mirror to check her make-up wherever she is. However, to improve it, she must first appropriate her own image in the reflection of the mirror, so as to check her work with her subtle beauty instruments. How can we recognise such a reflection as our own image? According to Husserl or Merleau-Ponty, the usual way of constituting the objects around us is what we call "the sketch donation". Let us examine if such a donation is compatible with mirrors.

2. The "sketch donation" phenomenology

When we are in motion, we can recognize the volumes of the objects we are unable to move. To do so, we work on the bypassing of the object. This process can be seen as a combination of positive isometrics, isometrics being transformations that keep distances as they are, and thus do not distort the transformed object. They are positive when they keep the internal sides and aspects of the object as they are. Within this combination of isometrics –mainly translations and rotations– that gives us access to the object we are aware of, rotation is essential to discover its hidden parts and to come to an adequate knowledge of it. Here, Husserl would talk of "sketch donations" to indicate that the perceptions of the object are always fragmented, and of "filling constitutions" to give meaning to the synthesis that leads us to a unified representation. In virtual reality, this is the way we create volumes: by scanning during motion capture operations, or by synthesis of 2D images. But through the mirror, nothing goes right! No possible filling: the reflected movement of my gaze yields sketches that do not lend themselves to a passive synthesis, but go on leading us to short-lived donations. In front of the mirror, we are inclined to recognise our own image, with only one exception: the *inversion* of laterality. Indeed, negative isometrics such as symmetry do not exist in nature except in the mirror and constitute a provocation to common sense, thus explaining our narcissistic fascination with it. The image that I see is a double of my face that faces me as in a face to face, and I could replace it by turning down or stepping back. This is the way we recognise our own images in mirrors.

[1] « The systematic approach to interaction design is characterized by a study of user tasks, existing interaction techniques, and characteristics of the user, environment, or system that might affect performance. In general, this approach is slow and methodical, with incremental improvements in performance rather than sudden leaps », Doug A. Bowman, Kruijff, LaViola, and P. Iaz. An introduction to 3D user interface design. 2001

3. The bypass of the "sketch donation"

When we are doing this, we are trying to turn by thought a 2D symmetry (negative isometric) to a rotation (positive isometric), which is geometrically impossible, except if we adapt to the illusion, and claim there is a left-right inversion. We thus reinforce Husserl's constitution with a virtual gesture that consists in being ourselves here at the same time as another there, in front of us.

Then, our own image in the mirror needs a *virtual* movement to be recognised as such, and it is so, regardless of geometry, which has at the same time prejudicial consequences for the one who is looking since he has to sacrifice the laterality of his movements and be prisoner of the reflection world. This gesture marks the start of a different kind of identification that no act donation will ever correspond to. Instead, we rather have sketch donations led by a fake who is just our own reflection, and this is not unsafe: as the 2D image of a 3D object cannot be understood without its prolongation in a 3D scene, its understanding depends only on its interpretation. If we call 3D we the space that enables us to make a sketch donation with real movements, we can imagine the virtual 3D space as the minimal space that permits the synthesis of virtual sketches.

4. Conclusion

Mirrors are far from being the user-friendly instruments that their generalised appropriation makes us tend to believe. On the contrary, the productivity they give comes with an extreme cognitive constraint, whose relaxation can only be obtained through the distortion of the most fundamental principles of geometrical topology. However, these strange instruments make us live a fascinating and fundamental experience that none of us would abandon, even though it can also be very alarming.

Our research and development of enaction tools and an enaction environment could cast doubt on the methodological un-thought of the "*convivial whole*" and lead to daring deep cognitive breaks. But is it worth it?

References

[1] Doug A. Bowman and C.A. Wingrave. Design and evaluation of menu systems for immersive virtual environments. In *Proc. IEEE VR'2001*, 2001.

[2] Doug A. Bowman, Kruijff, LaViola, and P. Iaz. An introduction to 3D user interface design. *Presence 10 (1)*, 2001.

[3] Doug A. Bowman, D. B. Johnson, and L. F. Hodges. Testbed evaluation of virtual environment interaction techniques. In *VRST '99: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 26–33, New York, NY, USA, 1999. ACM Press.

[4] Enactive'06, Third International Conference on Enactive Interfaces, Montpellier (France): November 20-21, 2006

[5] Jean-Christophe Bailly, *Le versant animal*, Bayard Centurion, 2007

[6] Jorge Luis Borges, *Œuvres complètes*, La Pléiade Gallimard, 1993

[7] Lewis Carroll, *Tout Alice*, Garnier-Flammarion, 1979

[8] Gilles Deleuze, *Différence et répétition*, Presses Universitaires de France, 1968

[9] Eugen Fink, *Le jeu comme symbole du monde*, Les éditions de minuit, 1966

[10] Edmund Husserl, Idées directrices pour une phénoménologie, Tel Gallimard, 1950

[11] Agnès Minazzoli, *La première ombre*, Les Editions de Minuit, 1990

[12] Friedrich Nietzsche, *Ainsi parlait Zarathoustra*, Gallimard - Folio / essais n° 8, deuxième partie : " L'enfant au miroir ", pages 108-109, Paris, 1985

[13] Andeï Tarkovski, *Zerkalo (Le miroir)*, film 106 min, couleur, avec Margarita Terekhova, Oleg Lankovski, Philip Lankovski et Ignat Daniltsev, URSS,1974

[14] Hermann Weyl, *Symétrie et mathématique moderne*, Champs Flammarion, 1964

Mirror-symmetry modulates perception dynamics

Salesse Robin*, Kostrubiec Viviane*, Zanone Pier-Giorgio*
*EA 3691 -Laboratoire Adaptation Perceptivo-Motrice et Apprentissage
UFR STAPS, Université de Toulouse Paul Sabatier
118, route de Narbonne, 31062 Toulouse Cedex 9, France
E-mail: salesse@cict.fr, kostrubi@cict.fr, zanone@cict.fr

Abstract

Motor and perceptual patterns are known to show equivalent properties. Recent studies on inter-limb and perceptual patterns cast some doubts on this claim. Traditionally, the stability and accuracy for motor relative phase is larger for mirror-symmetrical than for mirror-asymmetrical space-time relationships. Contrary to motor relative phase, perceptual relative phase is judged to be more stable and accurate for mirror-asymmetrical rather than mirror-symmetrical space-time relationship. In order to better understand this apparent inconsistency, we have a specific to-be judged perceptual relative phase defined as a pure time relationship. Participants were asked to reproduce the relative phase of to-be-judged stimuli using a psychophysical method of adjustment. Our results showed that, in accordance with the studies on interlimb patterns, the symmetrical time relation was reproduced with higher stability and accuracy than the asymmetrical one. This work suggests that, as far as the space-time properties of perceptual patterns are controlled, perceptual and motor relative phase display equivalent accuracy and stability properties.

1. Introduction

Numerous studies on perceptivo-motor patterns have provided a common picture on the properties of these two (perceptive and motor) systems [1,2]. However, recent studies on inter-limb and perceptual patterns cast some doubts on this equivalence. On the one hand, motor inter-limb patterns are classically captured trough space-time relations between horizontally moving limbs, as quantified by motor relative phase [3,4,5]. Traditionally, motor relative phase stability and accuracy is larger when limb movements produce mirror-symmetrical (i.e., symmetric oscillations with respect to the body midline, resulting from simultaneous flexion or extension of both wrists) as compared to mirror-asymmetrical space-time relationship (i.e., one limb moves towards the body midline, while the other moves away). On the other hand, perceptual pattern has been recently assessed through space-time relations between two visual targets moving horizontally, and quantified by perceptual relative phase [6]. Contrary to motor relative phase, perceptual relative phase was judged to be more stable and accurate for mirror-asymmetrical (visual stimuli represented by white dots on a black background are moving in the same direction, that is under an asymmetrical motion) than for mirror-symmetrical (visual stimuli are moving in opposite directions, that is under a symmetrical motion) space-time relationship. In order to better understand this apparent inconsistency, we asked the participants to judge perceptual relative phase specified using time relationship only (i.e. without movement).

2. Method

2.1. Participants

Ten right-handed adults (2 women and 8 men, M=22.7 years) volunteered for participation in the experiment. Participants were naive to the purpose of this experiment.

2.2. Apparatus and Materials

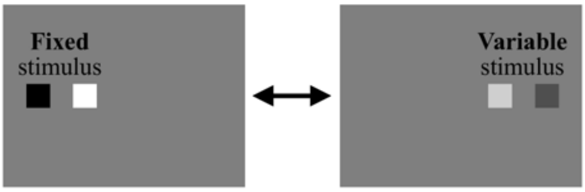
Participants were seated in a dark room. They viewed a computer screen and responded on a keyboard placed on the same table as the screen. The visual stimulus consisted in a pair of squares, displayed on a grey background. The luminance of each square oscillated from white to black, this variation being controlled by a quasi-harmonic periodic variable ($T = 1s$) corresponding to the grey level ([0 0 0] for black to [1 1 1] for white in RGB colors), and involving 144 grey levels per cycle.

Luminance difference between the two squares was controlled trough the relative phase between these two grey level variables. More precisely, each of both squares varied periodically in luminance intensity, the

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19^{th}-22^{nd}, 2007

relative phase corresponded to the difference between these two cycles.

On the screen, two visual stimuli could be displayed, a fixed stimulus on the left (the "model") and a variable stimulus on the right half of the screen (the "to be adjusted"). For fixed stimulus, thirteen fixed stimuli were used, characterized by one of the thirteen phase relationships from 0° to 180° by 15° steps. For variable stimulus, color relative phase could be decreased, or increased, over the whole 0°-360° range by 5° steps using the left or right arrow keys press, respectively.



Fixed and variable stimulus never appeared on the screen simultaneously. Instead, participants could switch from one to another using left or right shift keys press. This method was chosen to prevent participants from adjusting the variable stimulus using other strategies than the perception of the fixed stimulus color relative phase. Simultaneous press on F1 and F12 keys ended the current and started the next trial, separated by a gray screen during 8 seconds. Except the six above-mentioned keys ("left shift", "right shift", "F1", "F12", "left arrow" and "right arrow"), all the keys of the keyboard were inactivated. The whole procedure and visual display was controlled using the Matlab© Psychophysics Toolbox [7,8] on a customized 3GHz computer.

2.3. Task

According to a traditional psychophysics method of adjustment, [9] participants were instructed to try to adjust a pure time-defined relative phase of a variable visual stimulus until it looked like the relative phase of a model fixed stimulus. Each trial lasted until the relative phase adjustment was judged satisfactory.

2.4. Procedure

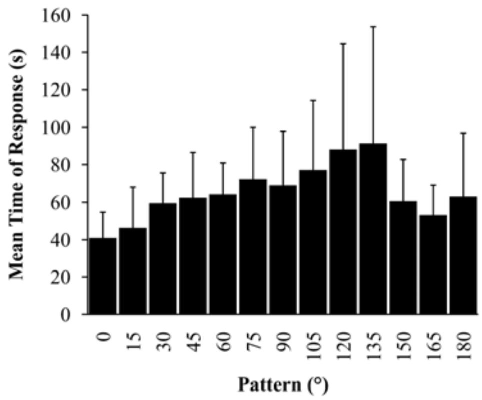
At the beginning of the experiment, participants performed four familiarization trials involving four relative phase values randomly chosen between 180° and 360°. Immediately thereafter, thirteen fixed stimulus were presented four times at random, providing 52 trials (4 trials x 13 relative phases). At the beginning of each trial, the initial value of the variable stimulus was shifted by ±45° apart from the fixed stimulus value, in order to keep constant the initial relative phase shift between the two stimuli. For each fixed stimulus, two trials started with -45° and two with a +45° relative phase shift. The fixed stimulus appeared on the left part of the screen. Participants were free to scrutinize it and/or to switch toward the variable stimulus and adjust it in reference with the model stimulus, as long and as many times as they want. Whereas the fixed stimulus color phase never changed during a trial, the variable stimulus could display every phase value within the range [0° – 360°] as a function of the participant's key presses, which in- or decreased the relative phase by steps of 5°. Although there was no explicit time constraint, participants were asked to terminate the trial as fast as possible.

2.5. Dependent variables

Four dependant variables were analyzed: 1) mean trial length, 2) mean constant error, 3) mean absolute error, and 4) the standard deviation of mean constant error. The trial length corresponded to the time used by the participant to achieve the adjustment task, which is the time delay between the beginning of the trial and the moment he decided to end it. This measure is thought to be an indicator of the adjustment difficulty [10]. To assess adjustment accuracy, the constant error was computed as the signed difference between the required relative phase and the relative phase produced by the subject at the end of each trial. The absolute error corresponded to the absolute value of the last-mentioned difference. The standard deviation of the constant error was used as a measure of the stability of the perceived pattern [11]. Each of the four dependent variables was analyzed using a variance analysis (ANOVA) 13 (pattern) with repeated measures. When necessary, post-hoc Newmans-Keuls analyses were carried out.

3. Results

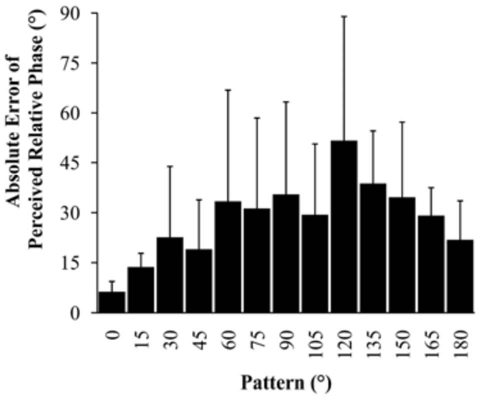
3.1. Trial duration



A 13 (pattern) repeated measures ANOVA carried out for trial duration revealed a main Pattern effect

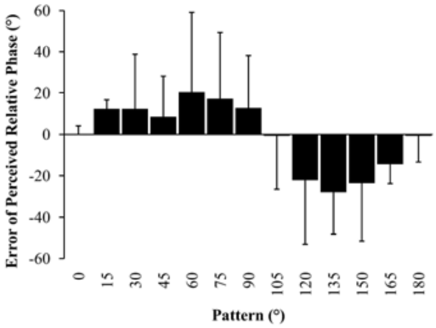
(F(12,108)=5.08, p<.001, η^2=.36). The distribution of trial duration as a function of patterns took an unimodal, asymmetrical shape, slightly left-skewed towards the in-phase pattern (0°), with a maximum centered around 135° pattern (65.2 s). The shape of this dataset suggests that patterns close to in-phase and anti-phase were adjusted faster than patterns located elsewhere. A subsequent post-hoc Newman-Keuls analysis indicated that trial duration was equivalent for patterns located near in-phase pattern (between 0° and 60°) and near anti-phase (between 150° and 180°). More specifically, the 135° pattern trial duration was significantly higher than that for patterns located closer to the distribution tails, between 0° and 60°, and between 150° and 180°. Trial duration for the 120° pattern was also higher than for patterns ranging between 0° and 15°, and for 165°. For the 105° pattern, trial duration was significantly higher than for patterns located between 0° and 15°. Finally, the 75° pattern trial duration was significantly higher than that of 0°.

3.2. Absolute error of relative phase



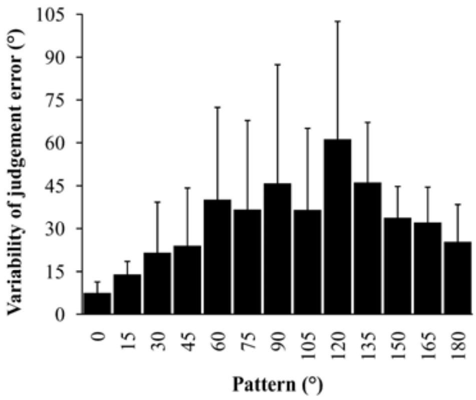
A 13 (Pattern) repeated measures ANOVA carried out on the absolute error revealed a significant Pattern main effect (F(12,108)=4.77, p<.001, η^2=.35). The distribution of absolute error exhibited a unimodal, asymmetrical shape skewed towards the in-phase pattern (0°), with a maximum at 120° (51.6°). This shape suggests that patterns located near in-phase and anti-phase were more accurately judged than others. Subsequent post-hoc Newman-Keuls test indicated that for 0° in-phase pattern, absolute error was lower than for patterns ranging between 60° and 165°. For 180° anti-phase pattern absolute error was lower than for 120° pattern. The 120° pattern revealed a significantly higher absolute error than patterns located between 0° and 45°, as well as the 180° pattern. Finally, absolute error of 135° pattern was significantly higher than those of 0° and 15°.

3.3. Constant error of relative phase



A 13 (Pattern) repeated measures ANOVA carried out on the constant error revealed a main effect (F(12,108)=4.88, p<.001, η^2=.35). Constant error distribution took a sine-like shape, characterized by a overshoot of relative phase for patterns ranged between 0° and 105° and an undershoot for patterns between 105° and 180°. Constant error peak was located at 60° pattern (error=+20.4°), and constant error valley at 135° pattern (error=-28°). For 0°, 105° and 180° patterns, constant error was null. Post-hoc Newman-Keuls test indicated that for 120° and 150° patterns constant error was lower than for patterns ranged between 15° and 90°, except for 120° with respect to 45° and for 165° with respect to 60°.

3.4. Standard deviation of the relative phase



A 13 (Pattern) repeated measures ANOVA carried out for constant error standard deviation revealed a main Pattern effect (F(12,108)=3.44, p<.001, η^2=.27). The standard deviation distribution showed a unimodal, asymmetrical shape, left-skewed towards the in-phase pattern (0°) with a maximum at 120° pattern (61.2°). This dataset shape suggests that constant error stability was higher for patterns located near in-phase and anti-phase than for patterns located elsewhere. Subsequent post-hoc Newman-Keuls test indicated that constant error standard deviation was equivalent for patterns located near 0° and 180°. For

120° pattern, standard deviation was reliably higher than for patterns ranging between 0° and 45° and for patterns between 165° and 180.

4. Discussion

In accordance with the studies on interlimb motor patterns, our results indicated that, a symmetrical time relation (0° of relative phase) was reproduced with higher stability and accuracy than an asymmetrical one (180° of relative phase), even in the absence of any information pertaining to movement. Stability and accuracy took the form of a classic U-inverted curve as a function of to-be-reproduced relative phases. The present work suggests that, as far as the space-time properties of perceptual patterns are controlled, perceptual and motor relative phase display equivalent accuracy and stability properties. It implies that perceptual dynamics depends on space-time symmetry properties rather than directional properties referred to an external frame of reference [12].

References

[1] Kelso, J.A.S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.

[2] Gibson, J.J. (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.

[3] Kelso, J.A.S. (1981). On the oscillatory basis of movement. *Bulletin of the Psychonomic Society*, *18*, 63.

[4] Kelso, J.A.S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*, *246*, R1000-R1004.

[5] Mechsner, F., Kerzel, D., Knoblich, G., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature*, *414*, 69-73.

[6] Bingham, G.P., Zaal, F.T., Shull, J.A., & Collins, D.R. (2001). The effect of frequency on the visual perception of relative phase and phase variability of two oscillating objects. *Experimental Brain Research*, *136*, 543-552.

[7] Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vision*; *10*; 433-436.

[8] Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*; *10*; 437-442.

[9] Fechner, G.T. (1860). *Elemente der Psychophysik*. Breitkof und Hartel, Leipzig. Reprinted Thoemmes Press, Bristol, 1999.

[10] Temprado, JJ., Zanone, P.G., Monno, A., & Laurent, M. (1999). Attentional load associated with performing and stabilizing preferred bimanual patterns. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1579-1594.

[11] Bingham, G.P., Schmidt, R.C., & Zaal, F.T. (1999). Visual perception of the relative phasing of human limb movements. *Perception & psychophysics*, *61*, 246-258.

[12] Swinnen, S.P. (2002). Intermanual coordination: from behavioural principles to neural-network interactions. *Nature Reviews Neuroscience*, *3*, 350-361.

Towards a Method of Eliciting Tacit and Embodied Knowledge for the Design of Multimodal Interfaces

Jack Sam Diane Gromala Lyn Bartram
Simon Fraser University, Canada
E-mail: jsama@sfu.ca, dgromala@sfu.ca, lyn@sfu.ca

Abstract

This paper presents the rationale for the research method we designed to elicit the embodied and tacit knowledge of race drivers. This knowledge will be used to create multi-modal racecar interfaces. The advancements in electric motor research allow automotive designers more flexibility in creating cars that are not restricted by large, conventional mechanical components. This research is meant to inspire exploration into conceptual driving interfaces that are made increasingly possible by the advent of electronic sensors and drive-by-wire technology. Our method consists of a questionnaire and a participatory design workshop which is intended to harness the race drivers' creativity by enacting their expertise in a simulated racing context. The drivers brainstorm, or ideate unconventional interfaces by focusing on innovative uses of sensory input/output, such as taste, smell or proprioception in order to more evenly distribute sensory and cognitive load during racing. This method is specifically designed to obtain unique ideas beyond conventional racecar design.

1. Introduction

We are studying the tacit and embodied knowledge in race drivers because we want to find out how useful information about traction can be gathered, interpreted and mediated, in order to design unconventional and innovative interfaces that allow for more direct control of a racecar. We extend and appropriate design researcher Liz Sanders' *MakeTools* [1] method by applying phenomenological and scientific perspectives specifically for race drivers. Our novel approach explores the deeper issues – that conventional focus groups can't reach – required to emancipate ourselves from the steering wheel/pedals paradigm. This method includes an apparatus, affording the experts a simulated spatial envelope that allows them to describe scenarios and tell stories using their entire bodies.

2. Novel Interfaces Matter

There is a significant disconnect between race driver experience and how vehicles are designed. Current approaches to racecar design focused only on efficiency and the standardization of drive-train components. The goal of the racing is to achieve fast lap times by continuously maintaining 100% usage of the tires' traction, however, there are no information displays for traction – like tachometres that indicate engine RPM – to communicate traction to the driver.

A newly developed racecar is a multimillion-dollar paradox. Each year, millions of dollars are invested into a new racecar. Formula 1 represents the pinnacle of automotive achievement and yet, its spirit of experimentation that once thrived in its rich history, is suppressed in favour of the homologation of car specifications to allow lower budgeted teams to remain competitive. These racecars are continuously being improved, but based on an old conventional archetype that assumes that the perfect way to interface with a car has been achieved.

Elite race drivers still only use sympathetic information about the racetrack from mechanical linkage-based control devices. They use vibrations from the steering wheel and foot pedals to find the traction necessary to achieve their extreme speeds. Entirely novel information display and control systems are possible – these may achieve more intimate relationships between the driver and the racecar. This sensitivity may in turn allow for the best decisions to be made about aspects of traction and can thus maximize the potential of the racecar.

By involving race drivers at the early stages of the design process, interfaces are likely to be designed outwards from the driver's perspective, rather than inwards from an arbitrary exterior shape. The long term benefits to challenging the status quo of racecar design is to set engaging, pioneering examples of breaking away from the conventional archetype of a car. This in turn may inform and inspire future designers to create a better vehicle for everyday life.

3. Importance of Embodied Knowledge

Race drivers embody and enact knowledge that can be harnessed to make interfaces that facilitate time-critical decisions. A racecar's four tires are its only means of contact with the racetrack so feedback to the driver about traction usage is one important factor in reducing lap times. As racecars are designed with increasingly more information and control mechanisms, the cognitive body must extend itself. Drivers learn to seamlessly articulate their prosthetic racecar through the information pathways that connect the body to the machine: the eyes to see the road, the ears to hear the engine, and the torso to feel traction.

According to Dourish's theories on embodiment, interpreting one's level of traction is similar to "using a stick to feel your way in the dark. You have the sense of exploring the ground in front of you (distal) while, in fact, what you are experiencing is a set of sensory impressions at the hand holding the stick (proximal)" [2]. This ability of a racecar driver to extend his/her body to include the racecar, in order to find the limits of traction, is what defines the driver as a skilled expert in operating vehicles.

Jens Rasmussen, on the topic of cognitive engineering, proposes the need to "analyse, formulate and make explicit the representations behind 'tacit knowledge'. The entire background of intuition giving meaning to explicit knowledge as formulated by skilled people during their professional interaction…has so far escaped the interest of most research programmes" [3].

Gary Klein offers his Recognition-Primed Decision model to examine decisions that are made in time-critical situations. This model leverages the knowledge-based of Rasmussen's Skills, Rules and Knowledge (SRK) classification. Race drivers have developed intuition that can help inform decisions where information is insufficient by breaking a larger problem into small familiar ones [4]. These cognitive processes are usually evoked unconsciously by sensory input. The speed of human affect is nearly instantaneous but the speed of consciousness and reflection is not. Therefore, a successful method would have to address race drivers' precognitive levels.

Henri Bergson's notion of time separates mathematical duration as segmented movements in space whereas real duration is a continuous subjective experience [5]. Ex-Formula 1 driver Jackie Stewart says, "[Speed] doesn't exist for me except when I am driving poorly. Then things seem to be coming at me quickly instead of passing in slow motion and I know I'm off form" [6]. Similarly, the late Ayrton Senna talked about his temporal sense in terms of living in the future. "I am able to get to a level where I am ahead of myself; maybe a fifth of a second, who knows? When my car goes into a corner, I am already at the apex" [6].

4. The Choice of Open-Wheel Race Drivers

We specifically targeted open-wheel (formula-style cars) race drivers who drive on road courses. Open-wheel drivers were targeted because their cars have greater contact between the driver and the environment than racecars that were modified street vehicles. These drivers are masters at interpreting a feedback loop between action and perception – the control of the car and its response in the environment – especially with their exposed cockpits. Along with this mastery comes latent needs and desires that concern issues such as cognitive load and anthropometry in their current racecar interfaces. We wish to harness this mastery and use it to create new interfaces that augment the racecar driver's goal of achieving better speed.

Race drivers are the ultimate usability testers for vehicular interfaces but their ideas are even more valuable. It is the passion of race drivers to seek their idiosyncratic idea of perfect synchronization with their racecar. Conversely, race drivers may have bottled-up frustration with the limitations of interfaces they have experienced, so they themselves can contribute to racecar design as cathartic expression.

5. Method Design

In order to employ a driver-centred approach to designing racecars, we designed a participatory workshop method based on Elizabeth Sander's Creativity-Based Research Tools model. The workshop uses physical sketching techniques as a means of brainstorming and documenting interface ideas onto a full-size mock racecar chassis. This abstraction chassis adds a context and a space restriction, essentially creating a simulation to try and capture the feeling of being in a real racecar for the participants. Sanders' method captures participants' unconscious actions by empowering them to impart their tacit knowledge and experience as brainstorm ideas and by getting them to imagine themselves in the future [1].

Upon recruitment, participants are given a set of questions to answer before participating in the workshop. This is based on Elizabeth Sanders' participatory design method. It tries to allow participants to communicate ideas that are not easily said in interviews or written in surveys. This method goes beyond observing what people do in their contextual environment by emphasizing what people make.

5.1. Recruitment

The workshops will be conducted in groups of 3. The participants are recruited race drivers who frequently race at River's Edge Road Course at Mission Raceway Park in Mission, BC, Canada. The president of the Sports Car Club of British Columbia has granted permission to solicit club members/race drivers as potential participants.

5.2. Questionnaire

The questionnaire is designed to both retrieve demographic data as well as to allow the drivers to reflect on what racing means to them in their life. This set of questions prime participants to leave their current racing situation behind so that when they arrive at the workshop, they are ready to speculate on their latent needs and desires.

5.3. Workshop Overview

Our adaptation of her method is designed to get race drivers to physically sketch, with paper and Velcro, interface ideas onto the abstraction of a chassis. This chassis situates the participant within the context of a racecar. Participants are encouraged to discuss their ideal interfaces and to not limit themselves to current racecar configurations. Two video cameras will capture the actions and discussions that may escape notice. The final results will manifest as a collection of unconventional interface sketches that were brainstormed by the experts themselves.

5.3.1. Apparatus

We created a mock racecar chassis for the race drivers to sit in during the workshop. The chassis serves to simulate an approximate envelop of movement and the affordances available when the driver is belted in tightly. Its dimensions are modeled after a 1979 Zink Formula Vee chassis. It is made from PVC pipes connected by rubber joints and covered with strips of Velcro (Figure 1).

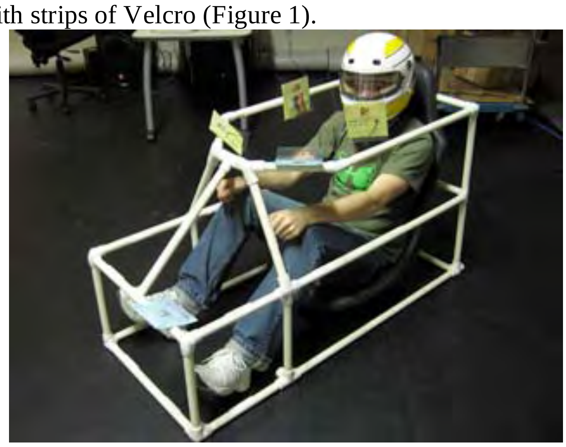


Figure 1. The first iteration of the mock chassis

To help document ideas, there will be three categories of components that can be quickly combined to express an idea. These are: a sticker representing a modality, a transparent plastic sheet with a printed image of a driving aspect and a sheet of coloured cardstock that allows for choosing either the reception of information or the control of a driving aspect. The sticker, plastic sheet and cardstock can be stuck together. Notes can be written on the card stock to further describe the participants' ideas. These will attach by Velcro onto the abstract chassis (Figure 2).

Figure 2. An example of an Interface Idea

5.3.2. Rules and Procedures

Upon arrival at the workshop, the participants will bring their questionnaires and their own helmets. The participants are given a verbal and written introduction to the workshop. They are instructed to brainstorm unconventional interface ideas and attach that idea by Velcro onto their preferred location on the cockpit. Each participant gets a turn to physically sketch within the cockpit while the other participants help on the outside. This abstract chassis serves to help situate the participant within the context of the racing car.

Participants are encouraged to discuss their ideal interfaces and not limit themselves to current racecar configurations. Their helmets and their bodies can also serve as a place to attach ideas. Participants are not allowed to denigrate the ideas of others for the sake of innovation. Each driver spends approximately 15 minutes inside the chassis.

5.4. Media Enriched Variations

After the pilot run of the workshop, the role of mirror neurons will be examined with the enrichment of media during the workshop. This media is introduced in two forms: video clips of race driving from the drivers perspective and playing a racing game with an electronic steering wheel and pedals set. The goal is to see if media helps to better situate the race drivers in the context their own racing experience [7].

5.5. Preliminary Testing

A sample preliminary test result shows a participant who chose to use the auditory sense as a means of controlling the acceleration of the vehicle (Figure 2). The participant further elaborates that the car should go faster according to how loud he screams. This interface idea was then attached to the helmet, indicating that the auditory sensor should be mounted on the helmet near the mouth (Figure 1). This test showed that a second, stronger iteration of the chassis was necessary to endure repeated use (Figure 3).

5.6. Debrief Session

After each driver has spent his/her session in the mock chassis, the chassis is photographed to document the location of each idea. Then each idea is removed and attached to a poster-sized 2-D representation of the chassis in their corresponding places (Figure 3). When all three drivers have spent a session inside the cockpit of the chassis, a debrief session will occur. The three poster-sized image of the chassis will be presented to all the drivers. This discussion allows the participants to clarify and/or elaborate on interesting ideas that emerged from the physical sketching sessions.

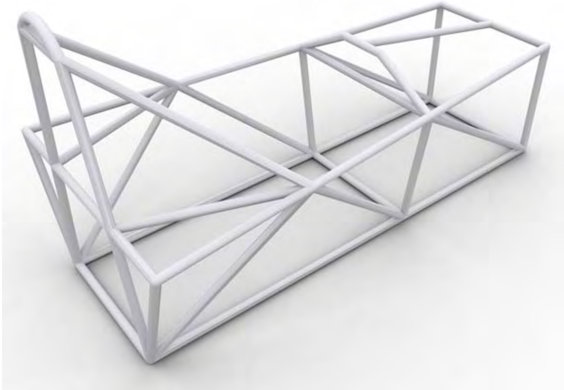


Figure 3. The second iteration of the mock chassis

6. Conclusion

The problem with eliciting expert knowledge is that this knowledge only resides in the body of the expert and is evoked through the contextual nature of the activity. Furthermore, upon evocation of this bodily knowledge, experts typically rest on the laurels of what they know and may have difficulty breaking away from that which they are most familiar with. In order to address these issues, we propose to meet experts halfway between what they know and what they need to abandon for the sake of innovation and creativity. By simulating the context-specific equipment in an abstract and incomplete way, we can evoke the experts' knowledge while not limiting them to the status quo.

Through our research, we will try to understand the intimate cognitive embodiment of race drivers' ability to perceive traction. We will elicit this tacit knowledge through participatory design methods. This research can then be applied to race driver decision-making models and address appropriate interfaces for this psychophysical state. Our position is that there are interfaces that have yet to be explored, especially considering that the senses of taste, smell, proprioception and explicit touch or haptics are neglected in current racecar designs.

This can be done through interfaces that facilitate the driver's body's ability to control the car while attention is used to make decisions and evaluate the body's actions. Designing racecars using a driver-centred approach could further the current state of integration between drivers and racecars. By transforming implicit embodied knowledge into explicit multimodal systems, we can redistribute and reallocate information according to race drivers' psychophysical abilities. Unconventional interfaces open up the possibility for better performance through the widening of and the opening of new information pathways between the driver and the car.

7. Future Work

15 race drivers have been scheduled to participate in the method described in this paper. The first group will act as a pilot while the following four workshops will be enriched with the media described in section 5.4. Once the interface ideas have been collected from the initial 15 participants, the most interesting interface ideas will be made into functional input devices for further testing with a racing simulator. There are also plans to extend our current method to include the study of participants' biopotentials in future studies.

The collection of interface ideas that result from our research will serve as a reference for inspiration that can be called up by anyone wishing to design any type of car. The workshop results will also inform the further refinement of this method for future use.

References

[1] Sanders, E.B.-N. "Virtuosos of the Experience Domain" In Proceedings of 2001 IDSA Education Conference, Massachusetts College of Art, Boston, MA, 2001.

[2] Dourish, P. Where the Action is: the Foundations of Embodied Interaction, MIT Press, Cambridge, MA, 2001.

[3] Goldstein, L.P., Anderson, H.B. and Olsen, S.E. Tasks, Errors and Mental Models, Taylor & Francis, London, UK, 1998.

[4] Klein, G. Sources of Power: How People Make Decisions, MIT Press, Cambridge, MA, 1998.

[5] Bergson, H. The Creative Mind, Greenwood Press, New York, NY, 1968.

[6] Bentley, R. Inner Speed Secrets: Mental Strategies to Maximize Your Racing Performance, Motorbooks, Osceola, WI, 2000.

[7] Gallagher, S. How the Body Shapes the Mind, Oxford University Press, New York, NY, 2006.

Business Models for Enactive Interfaces

Elisabetta Sani* Emanuele Ruffaldi* Haakon Faste* Massimo Bergamasco*
(*) Scuola Superiore S.Anna, PERCRO, Italy

E-mail: e.sani@sssup.it, e.ruffaldi@sssup.it, h.faste@sssup.it, bergamasco@sssup.it

Abstract

The definition of a Business Model is one of the fundamental steps in the creation of a new company, being strictly related to the types of product or services in the focus of the company itself. Among various types of technological products, Enactive Interfaces pose challenges in the definition of a good Business Model. The aim of this paper is to address these challenges and to provide an understanding of the peculiarities related to the commercialization of the Enactive Interfaces by analyzing a selection of companies in related fields.

1. Introduction

The concept of Enactive knowledge has brought new perspectives into the field of human computer interaction and Virtual Reality systems. While research teams are investigating the theory and creation of Enactive Interfaces, it is also interesting to analyze the possible business activities that these interfaces can introduce. The purpose of this study is to identify and assess current and potential business models relating to the creation and accessibility of Enactive Interfaces.

Despite widespread use in the industry, the concept of business models is often poorly defined. Some definitions address aspects related the architecture of the product, service and information flows [16]; others refer to the core logic for creating value in a company [13]; furthermore a series of authors introduce a financial element into their definitions [1], [14]. It is also worth mentioning the ongoing discussion of the difference between strategy and business models [15] even though many people use these terms interchangeably.

In general, the purpose of creating a model is to help understand, describe, or predict how things work in the real world by exploring a simplified representation of a particular entity. Therefore in the case of business models, the model shall be defined as a representation that helps an organization understand, describe and predict the 'activity of buying and selling goods and services'. Why is there interest in business models? Are business models useful? What are their purpose [10]?

The economic environment of today is competitive, complex [6], rapidly changing and characterised by an increasingly uncertainty [3] that makes business decisions complex and difficult. This difficulty increases when the level of the technology involved is even more sophisticated. As demonstrated by many studies [9], the business model concept can fill some of these gaps; it is one of the tools that can help tackle some aspects of the complexity by highlighting issues and pointing out the relationships between different variables, and can eventually gain an important position in managing uncertainty [5]. Business models can improve measuring, observing and comparing the business logic of a company [7]. Another advantage of the business model concept is to help foster innovation and increase readiness for the future through business model portfolios and simulation [2]; it may also be useful in the legal domain related to the patenting [4].

In the general context of Human Computer Interaction (HCI), Enactive Interfaces play a delicate role due to the fact that they interact more with the user's knowledge and sensorimotor system. Moreover, Enactive Interfaces can be analyzed in terms of their technology or application field and knowledge area. Some business models have been identified and applied to specific business fields such as web/digital cultural content, e-Business, and electronic markets [17], [16]. It has been found that some business models are much more common than others, and that some do, indeed, perform better than others.

A business model analysis approach has been selected and applied to a number of relevant companies in the fields of HCI and Virtual Reality (VR) systems, since clearly identifying Enactive related businesses is difficult. This approach allows the classification of companies into fields, and the identification of principal aspects that characterize their business models. After having analyzed the main aspects and challenges of Enactive Systems and the business model characteristics previously identified, a potential set of business models to be adopted for Enactive Systems has been proposed.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19^{th}-22^{nd}, 2007

As will be described in the following sections, the peculiarities of Enactive Systems affect the resulting business models. They require a strong focus on the user Human Perception Action loop, and—therefore—responsiveness and understanding to the needs of target users.

2. Analysis of Business Models

In addition to specific business model definitions and broad classification of such models, a number of authors provide us with business model ontologies [12], classifying business models with a certain number of common characteristics across a set of different categories. For simplicity, we have focused on an operative and systematic study [11]. Starting from a general definition it is possible to highlight business model components in key areas and their relationships. The entities of this ontology and their relationships are represented in Figure 1.

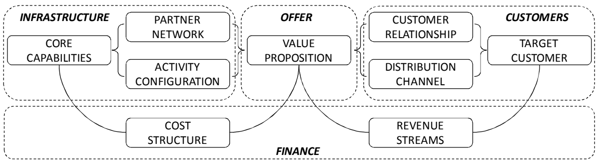


Figure 1. Business Model structure [12]

According to this model, the main aspect of a company is represented by its value proposition, strictly linked with its offer. This value is delivered to groups of customers through distribution channels. It is the strategic crafting of client acquisition channels that establishes lasting and trusting relationships with the customer. In most companies, these activities—and particularly understanding target customers and defining the customer relationship—are performed by marketing staff. Once having defined its core capabilities and infrastructure, the company generates value through the arrangement of resources and activities supported by partner relationships with other entities. Finally, the financial aspects of the company are supported by its cost structure and revenue streams.

3. Business Model Analysis in HCI and VR

The detailed structural analysis of a business model based on its value proposition can be used for the classification of business related to Enactive Interfaces. For this study, the starting point has been businesses focusing on human-computer interaction with an emphasis on Enactive Interface technology and design. Because many different technologies are involved in the creation of multimodal interfaces, companies have been selected to cover areas spanning hardware, software and servicing. The list has been obtained by focusing on companies with a strong position in the market or a well known history of success.

In the area of hardware we have selected producers and distributors of haptic interfaces, interactive interfaces and more traditional HCI devices. The companies representative of the software aspects are those related to software tools for managing advanced interfaces, or supporting human interaction based on awareness and knowledge. Finally, we have selected companies in the area of simulation and training, including servicing for such systems (Figure 2).

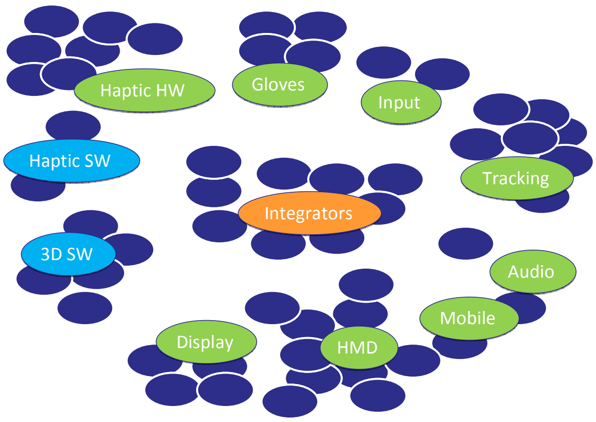


Figure 2. A display of the selected companies (blue circles) grouped by product area

The above empirical selection of 55 companies can be analyzed to understand possible business models according to two dimensions: involved assets and rights being sold. The first dimension distinguishes among four important assets: physical, financial, intangible and human. The second dimension classifies a business as creator, distributor, landlord or broker.

By combining these two dimensions it is possible to classify businesses according to sixteen different business models. Among these sixteen only few of them characterise the HCI-VR businesses; Figure 3 shows the distribution of the selected companies.

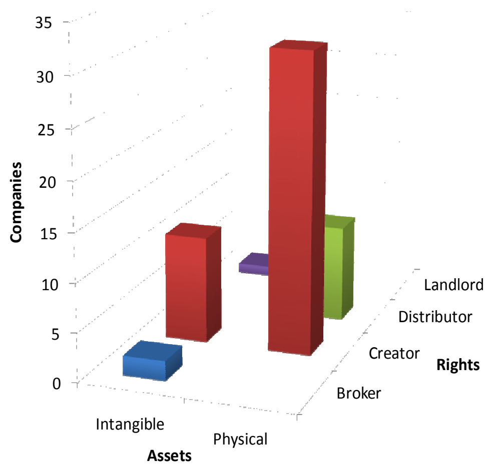


Figure 3. Distribution of Business Model types

The most recurrent role is represented by physical creator including companies that produce devices; this kind of business is strictly linked with physical distributors. From a general overview it is possible to note that the absence of some business models in this representation may indicate an opportunity to be taken into account, for example the areas of human training and exchange that have not yet received much coverage.

Since this analysis allows the classification of any type of business, the analysis has been focused to deal with key features and their relationship. In particular we adopt the business ontology description presented above [12]. The detailed analysis starts by comparing specific companies in different areas of HCI-VR.

In the field of haptic interface production three major actors are *Sensable*, *ForceDimension* and *Novint*. There is an important difference in the way *Novint* poses itself because its target is gamers. It is also worth mentioning that *ForceDimension* is performing a form of outsourcing providing the core technology for *Novint* devices. Figure 4 shows the analysis outlined in Figure 1 applied to Novint.

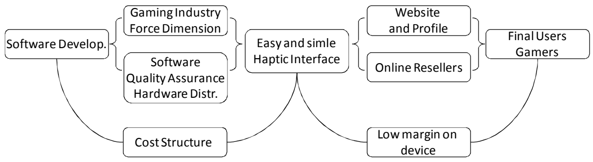


Figure 4. Structure of the Business Model of Novint

Once the ontology is applied to all of the companies in the sample, it is possible to construct an aggregate business model diagram showing all of the approaches and their different aspects. In the majority of cases the value proposition is a specific device or software, although there is a wide group of companies that provide consulting and complete VR systems. As expected, the target customers are mostly represented by industries and research centers, although a few companies are consumer oriented. Such customers are reached mostly by resellers, or with subsidiaries that sometimes provide demonstration centers. The value configuration is mostly based on support service for existing hardware, integration services for a better use of the devices, and some focus on device design. The core capability of a company is generally related to a specific technology. Finally, the partner network of such companies is characterized: for producers by companies that provide complementary technologies, or for integrators by the providers of core technology.

The Business Model ontology analysis applied to the selected companies allows the major structures of these companies to be identified, but is unable to capture dynamic aspects of the evolution from an idea to a product. We have therefore identified two main categories for capturing this evolution. The first indicates the status of an idea's maturity; in particular we have selected four main stages: idea, prototype, product and mass market. While this classification is common for general products that become mainstream, for HCI and VR it is more difficult to advance from one stage to the next. The second dimension captures the place in which such transformation takes place. Four principal areas have been identified: the University where many VR products originate; companies managing engineering or production of products; System Integrators, companies that assemble VR systems; and external producers that provide outsourcing or licensing companies.

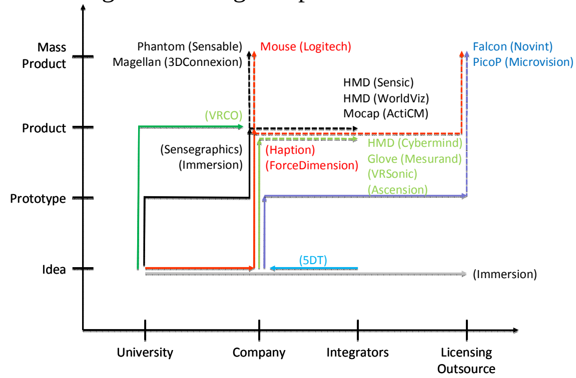


Figure 5. Description of evolution from idea to products along Product Paths

Given these two categories we have identified some *product paths* that show the relationship between an original idea and the places in which it becomes stronger. Figure 5 shows these main steps and examples of identified products and companies that proceeded along these paths. Different colours correspond to different steps, while the dashed lines correspond to possible directions for a path. In the case of Haptic Interfaces, some companies begun by developing prototypes in Universities, such as *Sensable*, or preliminary ideas. In the case of software it is common for complete products to be developed in a University that are later sold to a company for distribution. The grey line from idea to licensing represents the case of patent licensing.

4. Toward Enactive Systems

The creation of a business around Enactive Systems presents challenges similar to those of HCI, with some more specific aspects introduced by the characteristics of such systems.

The multimodal component of an Enactive Interface makes the interface more complex to manage, yet at the same time the goal of Enactive Systems is to create interfaces that are simple to use and intuitive.

Intuitiveness and usability are two challenging aspects that have been extensively addressed by HCI systems because of the large target market. In the case of VR devices, however, the usability aspect is only one of the dominant parameters for device quality because the target customers for such emerging technologies are often highly specialized and are willing to invest in learning new technologies despite the costs involved. As the target market for such devices grows, the importance of usability, brand identity and human-centered design activities becomes an increasingly important element of a product's success. In the field of VR, many companies have pursued a strategy of device production or technology development focus, provided that they have a partner network allowing them to tackle difficult integration problems and software aspects. In some cases this type of company has adopted a more vertical approach, producing the relevant software and providing customers with a complete solution. This requires the company to be prepared to offer the support needed to service and respond to evolving user needs.

In contrast to complicated VR installations, many Enactive Systems tend to be more simple and easier to use, being in many ways more similar to a traditional HCI device. For this reason, it is more probable to adopt a Business Model in which the complete system, including hardware and software, is managed and provided to the final user. A central aspect of an Enactive System is the strong relationship with human perception and understanding. This relationship is not limited to specific sensorial fields, as in HCI or VR systems, but involves multiple senses. The integration of such aspects for a wide market product requires a strong effort for the validation of the product and its final design. For companies with a heavy engineering focus, business and human factors experts should be involved early in the design process to develop viable products that respond to real-world need. Design research strategies can be useful in this process [8].

A viable solution to the complexity of Enactive Systems is licensing, a means that has been applied in many hardware and software cases. A company decides to license its technology when it is not able to produce it by itself, or when integration activities are beyond its core expertise. Such was the case during the initial phase of *Immersion*, when force feedback joystick technology was licensed to a third party for production. In its current stage, *Immersion* produces high-end devices for research and high-tech companies while licensing its force feedback technology to console companies like *Sony* and *Microsoft*. A similar approach has been followed by *DDD*, a company specialized in glass-free stereographic displays, who initially licensed its technology to *Shape3D* and later followed with the direct production of advanced displays.

In the Business Model ontology, the value proposition is not the only core aspect to be analyzed. The partner network allows problems concerning technology integration to be managed while at the same time exploiting the technology in different application fields. In the area of HCI, the success of peripherals depends on support from the videogame industry - in particular if such devices provide new gaming possibilities.

5. Main Conclusions

The identification of future business models and their effectiveness has some limitations, but the results of this analysis can provide direction and understanding to the development of more effective businesses concerning Enactive Interface technology. Hopefully, the application of this knowledge will lead to the greater dissemination and use of Enactive Interfaces in the marketplace.

References

[1] Afuah, A., and Tucci, C. "Internet Business Models and Strategies: Text and Cases". McGraw-Hill Higher Education, 2000.

[2 Allen, P. "A complex systems approach to learning in adaptive networks". International Journal of Innovation Management 5, 2 (2001), 149–180.

[3] Andriani, P. "Diversity, knowledge and complexity theory: dome introductory issues". International Journal of Innovation Management 5, 2 (2001), 257–274.

[4] Angwin, J. "Business method' patents, key to priceline, draw growing protest". Wall Street Journal - Eastern Edition (2000).

[5] Courtney, H., Kirkland, J., and Viguerie, P. "Strategy under uncertainty". Harvard Business Review 75, 6 (1997), 67–79.

[6] Hodgson, G. "Capitalism, complexity, and inequality". Journal of Economic Issues 37, 2 (2003), 471–479.

[7] Kaplan, R., and Norton, D. "The balanced scorecard. measures that drive performance". Harvard Business Review (1992), 71.

[8] Laurel, B. "Design Research: methods and perspectives". MIT Press 2003.

[9] Linder, J., and Cantrell, S. "Changing business models: Surveying the landscape". Accenture Institute for Strategic Change (2000).

[10] Magretta, J. "Why Business Models Matter?" Harvard Business School Pub, 2002.

[11] Malone, T., Weill, P., Lai, R., D'Urso, V., Herman, G., Apel, T., and Woerner, S. "Do some business models perform better than others? a study of the 1000 largest us firms". Tech. rep., MIT Sloan School, Cambridge, MA, 2006. 4615-06. MIT SeeIt Project.

[12] Osterwalder, A. "The Business Model Ontology: A Proposition In A Design Science Approach". PhD thesis, l'Ecole des HEC de l'Université de Lausanne, 2004.

[13] Peterovic, O., and Kittl, C. "Developing business models for ebusiness". In International Conference Commerce Conference (2003).

[14] Rappa, M. "Managing the digital enterprise-business models on the Web". North Carolina State University (2001).

[15] Schein, J. "The role of the business model and strategy for business". Tech. rep., Mindsheet Ltd, 2004.

[16] Timmers, P. "Business models for electronic markets". Electronic Markets 8, 2 (1998), 3–8.

[17] Wall, G. "Business model issues in the development of digital cultural content". First Monday, Peer-Reviewed Online Journal (2006).

Motion Sickness and Postural Motion Affected by Stance Width

L. Scibora* S. Villard* T. A. Stoffregen*
(*)Affordance Perception-Action Laboratory, University of Minnesota, USA
E-mail: scibo002@umn.edu, svillard@umn.edu, tas@umn.edu

Abstract

We exposed participants to large-field anterior-posterior visual oscillations in a moving room while varying the width of side-by-side stance. Participants stood with their heels pressed together (0 cm), heel midlines 17 cm apart, or 30 cm apart. Room motion was a sum of sines between 0.1 and 0.4 Hz, with maximum amplitude of 1.8 cm. The magnitude of sway in the mediolateral axis was reduced in the 30 cm condition, relative to the 17 cm and 0 cm conditions. The incidence of motion sickness was reduced in the 30 cm condition, relative to the 0 cm condition. Head and torso movement differed between the Sick and Well groups prior to the onset of motion sickness. These findings bolster the notion that posture is linked to motion sickness. In addition, the results support the hypothesis that motion sickness incidence can be influenced by variations in stance width.

1. Introduction

The malaise of motion sickness induces physiological distress in as much as one third of the population [5], in spite of the fact that it is a normal response to the perception of motion. Motion sickness, *aka cybersickness*, is frequently reported by users of virtual environments (e.g. fixed base flight simulators) and entertainment systems (e.g. wide-screen films and video games). It is *visually induced motion sickness* that occurs with visual motion, inherent to virtual environments, relative to the illuminated environment. Thus, understanding the causality of this sickness should be investigated concomitant to the development of virtual technology with goals of predicting and preventing its occurrence.

Many theories of motion sickness, such as the sensory conflict theory, have focused on the notion of intersensory conflict. That the development of motion sickness is dependent upon a discrepancy between expected patterns of sensory stimulation and the current pattern of perceptual stimulation makes intuitive sense. This theory fails when one considers the virtual impossibility of a scientist to know an individual's history of sensory stimulation and then predict such a conflict based on the current stimulation.

Another theory, the postural instability theory of motion sickness, states that visual motion environments could lead to instabilities in the control of movement, in general, and of bodily posture and orientation in particular [6]. This theory proposes that motion sickness will follow the development of such instabilities in postural control, and that motion sickness would occur only among persons who exhibited postural instability.

We have identified differences in postural activity (prior to the onset of subjective symptoms of motion sickness) in previous studies, between participants who later became motion sick and those that did not (e.g., [1, 8 & 10]). In these studies, a moving room was utilized to produce oscillations of an optical display along the body's anterior-posterior (AP) axis. The AP visual oscillations induced changes in standing body sway, followed by motion sickness.

One factor known to influence the magnitude of postural sway in the body's mediolateral (ML) axis is stance width (the distance between the feet in side-by-side stance). That is, increasing the width of stance tends to decrease the magnitude of ML sway. During unperturbed side-by-side stance, spontaneous body sway in the ML axis is strongly influenced by the distance between the feet. When the feet are close together there is a great deal of ML sway, and when the feet are far apart ML sway is reduced (e.g., [2, 4 & 11]).

Perturbed stance may differ from unperturbed stance, particularly in situations that may be challenging to stance. Thus, manipulations known to produce a reduction in ML sway might reduce the incidence of motion sickness.

The present study aimed to evaluate the influence of stance width on postural activity, incidence of motion sickness, and tendency of changes in sway (during exposure to visual motion) to precede motion sickness.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19^{th}-22^{nd}, 2007

2. Methods

2.1 Participants

Fifty-six students from the University of Minnesota participated in the study. There were 22 males and 34 females with a mean age of 20 years, mean height of 170 cm and mean weight of 68 kg. All participants had corrected to normal or near normal vision. In addition, all participants stated they were in good health and did not have a history of dizziness, falling or vestibular disorders. In preparation for the study, participants followed instructions to not eat four hours prior to their participation.

2.2 Apparatus

Optical motion was created by a moving room (e.g., [5]), which consisted of a three-sided cubical frame, 2.44 m on a side, mounted on wheels and moved in one axis on rails. Movement of the room was powered by a motor under computer control. Participants stood on a force platform in the moving room. The experimental stimulus consisted of room oscillation using a sum-of-sines function comprising ten sine waves in the range 0.02 – 0.31 Hz. The maximum peak-to-peak amplitude of room oscillation was 1.8 cm.

2.3 Procedure

To assess an initial level of symptoms and to ensure familiarity with motion sickness symptomatology, participants completed a simulator sickness questionnaire, or *SSQ* [3]. Once completed, participants entered the moving room and stood on a force plate. In each condition, the feet were rotated outward to an angle of 10°. The heels were pressed together (the 0 cm condition), or their midlines were separated by 17 cm or 30 cm as designated by marked lines on the platform. The sequence of trials consisted of two 60 s trials absent of imposed motion followed by four 600 s trials of imposed motion using the sum of sines stimulus, and concluded with a repeat of the initial two 60 s trials.

Participants were instructed to discontinue the experiment immediately upon any symptoms. Following discontinuation or completion of all the trials, participants completed the SSQ a second time (post-exposure).

3. Results and Discussion

Nineteen participants stood with their heels pressed together (0 cm), 20 stood with their heels 17 cm apart and 17 stood with their heels 30 cm apart. We found a significant effect of stance on the incidence of motion sickness between the 0 cm and 30 cm conditions (Table 1). The 63% sickness incidence in the 0 cm condition (12 of 19 participants) was significantly greater than that of the 23% incidence in the 30 cm condition (4 of 17), chi-square = 0.017, $p < .05$, indicating the wide stance reduced the incidence of visually induced motion sickness.

Table 1. Stance width and incidence of motion sickness.

	N - Total	N - Sick	Sick Incidence
0 cm	19	12	63%
17 cm	20	9	45%
30 cm	17	4	23%

We compared postural motion during the sum-of-sines trials of the three conditions. It was expected that as stance width increased the variability of motion in the ML axis would decrease, and that a narrow stance width would result in increased variability of motion in the ML axis. We found a significant Condition X Axis interaction in the variability of torso movement, $F(2,49) = 4.175$, $p < .05$ (Figure 1). Mediolateral variability was reduced in the 30 cm condition, relative to the 17 cm and 0 cm conditions. Thus, the difference in AP variability and ML variability decreased as stance width increased.

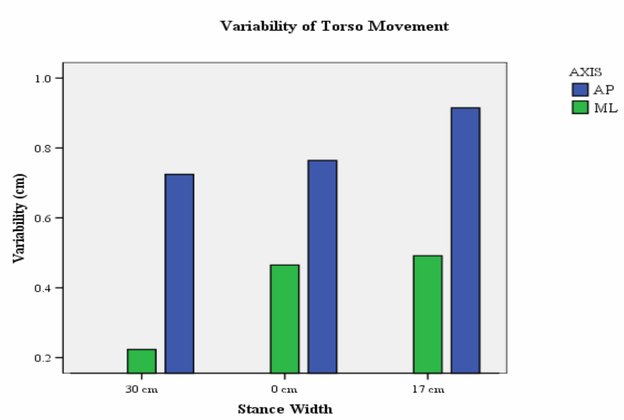


Figure 1. Condition X Axis interaction of the variability of torso movement.

In examining postural sway over time, we took three two-minute windows of time during exposure to the moving room. Windows from the beginning, middle and end of exposure were chosen based on the mean duration of exposure of participants in the Sick group. We found a significant effect of torso velocity between those participants who became sick (Sick group) and those who did not become sick (Well group), $F(2,98) = 3.436$, $p < .05$ (Figure 2). The Sick participants exhibited greater velocity of torso movement over time than the Well participants. In support of the postural instability hypothesis, greater movement was evident among Sick participants versus Well participants.

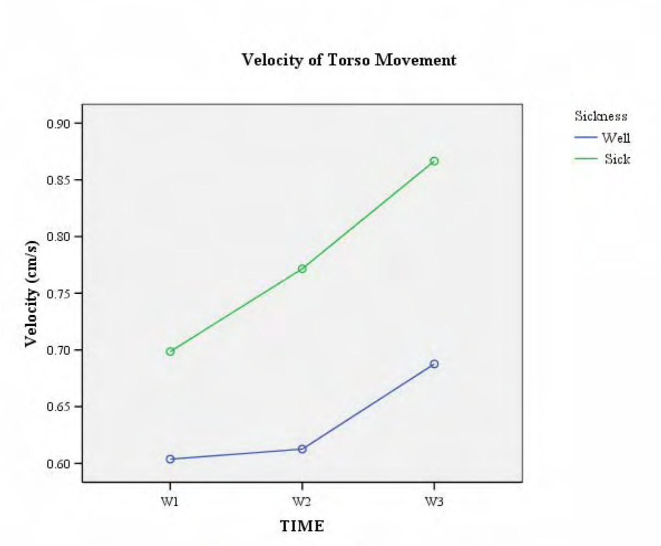


Figure 2. Group X Time interaction of the velocity of torso movement.

The results confirm that visually induced motion sickness is preceded by changes in standing body sway. The significant effect of stance width on motion sickness incidence indicates that a wider stance may reduce the incidence of motion sickness. This finding may have practical value in terms of instructions to motion sickness susceptible individuals in provocative environments. We studied perturbed stance width on postural sway in which the difference in variability between AP and ML sway decreased as stance width increased. This finding follows previous posturography data on unperturbed stance of ML sway decreasing with increasing stance width. Beyond these findings, the results may have implications for theories of postural control, in general.

References

[1] Bonnet, C. T., Faugloire, E. M., Riley, M. A., Bardy, B. G., & Stoffregen, T. A. (2006). Motion sickness preceded by unstable displacements of the center of pressure. *Human Movement Science*, *25*, 800-820, Champaign IL.

[2] Day, B.L., Steiger, M.J., Thompson, P.D., & Marsen, C.D. (1993). Effects of vision and stance width on human body motion when standing: implications for afferent control of lateral sway. *Journal of Physiology*, *469*, 479-399.

[3] Kennedy, R. S., & Lane, N. E. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology, 3,* 203-220.

[4] Kirby, R.L., Price, N. A., & MacLeod, D.A. (1987). The influence of foot position on standing balance. *Journal of Biomechanics*, *20*, 423-427.

[5] Lishman, J. R. & Lee, D. N. (1973). The autonomy of visual kineasethesis. *Perception*, *2*, 287-294.

[6] Murray, J. B. (1997). Psychophysiological aspects of motion sickness. *Perceptual & Motor Skills, 85*(3 Pt 2), 1163-1167.

[7] Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. *Ecological Psychology*, *3,* 195-240, Mahwah NJ, 1991.

[8] Smart, L. J., Stoffregen, T. A., & Bardy, B. G. (2002).Visually-induced motion sickness predicted by postural instability. *Human Factors*, *44*, 451-465, San Diego CA.

[9] Stoffregen, T. A., Hove, P., Bardy, B. G., Riley, M. A., Bonnet, C. T. (2007). Postural stabilization of perceptual but not cognitive performance. *Journal of Motor Behavior*, *39*, 126-138, Champaign IL.

[10] Stoffregen, T. A., & Smart, L. J. (1998). Postural instability precedes motion sickness. *Brain Research Bulletin*, *47*, 437-448, London.

[11] Tarantola, J., Nardone, A., Tacclini, E., & Schieppati., M. (1997). Human stance stability improves with the repetition of the task: effect of foot position and visual condition. *Neuroscience Letters*, *228*, 75–78.

Can haptic maps contribute to spatial knowledge of blind sailors ?

Mathieu Simonnet* Stéphane Vieilledent * Jean-Yves Guinard * Jacques Tisseau *
(*) European Center For Virtual Reality (CERV), LYSIC, EA3883, France
E-mails: simonnet@enib.fr, stephane.vieilledent@univ-brest.fr, Jean-yves.guinard@free.fr, tisseau@enib.fr

Abstract

In this preliminary study, we compared the capability of a blind sailor to access geographical information needed to navigate via an haptic device and via a tactile map. We assessed this spatial knowledge in an egocentered, an allocentered and a combined frame of reference. The subject first explored haptic or tactile maps before answering a series of questions in order to locate 6 salient objects within each map. Then, we used the triangulation technique to obtain easily scoreable physical representations of these cognitive locations. Basically, our results showed no difference between haptic and tactile condition even if slight differences were observed between the frames of reference. We suggest that the subject took great advantage of the haptic map because its sequential and dynamic features implied to focus on learning and memorizing the movement patterns rather than directly touching the global layout with reduced movements as it is the case when using a tactile map.

1. Introduction

In the present paper we define spatial knowledge as the capability to locate objects in the environment. When a human subject tries to determine his own position with respect to salient points, his entire body is also considered as an object of the environment. In particular, during navigation, ie a displacement of the entire body, the subject has both to gain knowledge of the position of his starting and final points and to update his current position.
Spatial knowledge is based on two kinds of information. The egocentered frame of reference refers to the individual point of view in a route perspective whereas the allocentered one implies an bird's eye view or maplike perspective [10].
All these pieces of information contribute to the constitution of cognitive maps, mental spatial representations, supposed to reproduce the spatial characteristics of the physical world. However since no isomorphism may exist between cognitive maps and the actual environment, it is often convenient for a human subject to use geographic maps. They provide the subject with global and allocentered information.
From a functional point of view, neither the actual displacement in the environment, nor the geographic maps used in isolation can lead to the building of a situated cognitive map that may improve spatial efficiency during navigation. Indeed, sensori-motor and symbolic information have to be merged in coherent action and mental spaces to identify common landmarks in the physical world and in the geographic map. Since vision is predominantly involved in this process, identifying the common landmarks remains a major difficulty for blind people during navigation.
In a triangle completion task during which subjects were accompanied along the two first legs of a triangle and then asked to walk back alone to their starting point, Loomis et al. [11] showed that blindfolded sighted, adventitiously and congenitally blind subjects obtained similar results. This result is in line with the conclusion of Rieser et al. [13] who previously found no differences between the same categories of subjects when they had to point manually toward their starting point after following a locomotor path. More recently, Gentaz and Gaunet [3] show that the inference process of a spatial location of a point is predominantly based on the manual movements. Nevertheless the absence of difference may be explained by the fact that the previous tasks do not require distant information. Indeed, the "difference theory" shows that blind people do not suffer of a lack of spatial reasoning but rather of difficulties in perceiving distant information [2]. Geographic maps can potentially minimize this problem by providing blind people with distant information that becomes available within their manipulatory space. Espinosa et al. [1] argue that the combination of direct experience and tactile maps constitutes a useful procedure which should be used by the blind people and Jacobson [6] verified the importance of tactile maps for helping blind and visually im-

paired people to form impressions of their surrounding space. Moreover, Rossano [14] emphasizes the importance of alignment in blind subjects' use of tactile maps. Usually maps are watched or touched with placing the north in front of the subject. However, during actual displacements in their environment the individuals never stay facing the north. This may lead to errors.

In this respect, virtual reality constitute valuable tools to provide blind people with a naturalistic and intuitive interface dedicated to the development of spatial knowledge. In particular, they can take advantage of haptic maps when getting the cartographic information via a computer controlled, motorized device held in the hand. Such a device produces force feed-backs when the user touches a virtual object. In some circumstances haptics can substitute for other sensory modalities like vision [12].

For example, Jansson [8] attempted to enable blind people to touch virtual geographical environments with an haptic mouse and with a Phantom Omni device, but the benefits of these new devices do not show real improvement. Later Jacobson [7] et al. use a force-feedback mouse and auditory labels or directions to give a mixed modal interface that allows more comprehensive feedback [4].

Here, in our preliminary study, and before the computer simulation of the displacement in its actual and virtual environments, we compare the capability of a blind sailor to access geographical information he needed to navigate via an haptic device and via a tactile map. We assessed this spatial knowledge in an *egocentered*, an *allocentered* and a *combined* frame of reference.

2. Method

2.1 Task and procedure

This preliminary experiment implied a congenitally blind thirty six years old male. During a first phase referred to as the *exploration phase*, the subject explored either a tactile or a haptic map. Then he answered 3 batteries of 18 questions during the *question phase*.

2.1.1. Exploration phase

Whereas the subject explored the tactile map using his two hands, he explored the haptic map with the Phantom device held in one hand only. These two maps of 30 cm by 40 cm contained a little part of land, a large part of sea and six salient objects.

On the tactile map, the sea was represented in plastic and the land was in sand mixed with paint. The salient objects were 6 stickers in different geometric shapes (e.g. "triangle", "rectangle", "circle") .

The haptic map came from SeaTouch, a JAVA application developed in our laboratory for navigation training of blind sailors. This software uses the classic OpenHaptics Academic Edition Toolkit and the Haptik library 1.0 final to interface with the Phantom Omni device. The contacts with geographical objects are rendered from a JAVA3D representation of the map and environment. Like a computer screen, this map stands in the vertical plane and implies that the north was at the top and the south is at the bottom. The rendering of the sea was soft and a sound of waves was played when the subject touched it. The rendering of the earth was rough and three centimeters higher than the surface of the sea. A sound of land birds was played when there was a contact with the land. Between the land and the sea, the coastline, as a vertical cliff, could be felt and followed with the sounds of sea birds. The salient objects were materialized by a spring effect when the haptic cursor entered in contact with them. Then a vocal synthesis announced the name of the objects (e.g. "rock", "penguin" or "buoy"). The exploration phase stops when the subject is confident to know the objects layout.

2.1.2. Question Phase

Among the 3 types of questions the two first one imply responses given either in an *egocentered* or an *allocentered* frames of reference. Nevertheless, in the third battery, questions required an answer combining the two frames of reference. This condition is referred as the *combined* one. In each case, each direction estimate of a salient object was given from three other objects.

In the *egocentered* frame of reference, the subject performed a pointing task from his own point of view to answer the following kind of question :
"From the penguin, could you point to the rocks ?" Here, we gathered the direction estimates (e.g. 32° on the left) thanks to a particular protractor.

In the *allocentered* frame of reference, the subject was asked to assess the directions with vocal answers. Here the subject was supposed to verbally respond in degrees from 0 to 359 (e.g. 270° for the west) to the following kind of question : "From the the penguin, where is the rock located ?"

In the *combined* frame of reference, our goal was to access to the situated cognitive map of the subject. In this case, the subject had to answer the following kind of question with the protractor :
"You are positioned at the penguin and facing at the rock, where is the buoy ?"
For example, the subject told us the point "penguin" was 45 cardinal degrees from the point "rock" in allocentric questions. Then he imagined he was at the "penguin" facing the "rock" and estimated the "buoy" at 36 degrees on the right with the specific protractor. Consequently,

we ruled off a 91 cardinal degrees oriented line from the "penguin" to the "buoy". Thus, we merged egocentric and allocentric responses.

2.2. Data reduction

In the present study, we used triangulation technique to obtain easily scoreable physical representations of cognitive maps. This method was originally adapted by Hardwick et al. [5] from the more familiar triangulation method used in navigation to determine the position of a ship. Typically, respondents estimated the distance and direction to a location from three or more places. The resulting vectors could be drawn and where the lines cross, a triangle of error could be outlined whose mean center was taken as the cognitive location of a place [9].
We measured i) the average euclidean distance between the actual objects displayed on the haptic or tactile maps and the cognitive location of their counterpart in the mental space and ii) the average area of the error triangles. We compared the different maps used during the exploration phase (*haptic* vs *tactile*) and the frames of reference (*egocentered*, *allocentered* or *combined*) within which the subject answered the questions by means of a repeated measurement analysis of variance (ANOVA).

3. Results

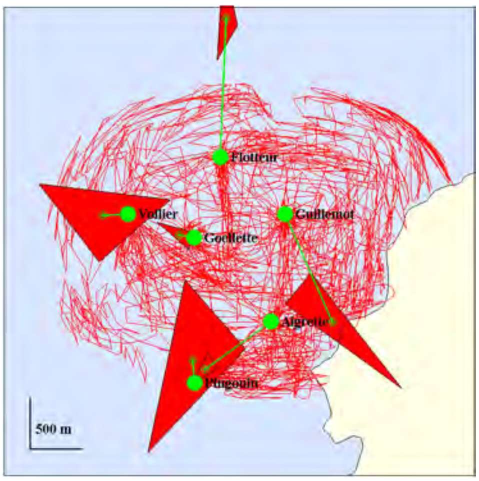


Figure 1: Error triangles built with projective convergence technique after haptic exploration and egocentered questions. The actual locations of the salient objects are displayed with large circles whereas their cognitive counterparts are depicted with small circles. The curved line is the trajectory of the haptic cursor during the exploration phase.

On a qualitative point of view (Figure 1), the haptic exploration of the virtual space shows that the subject clearly distinguished the sea and the land since his cursor never went beyond the coast line. Moreover, his exploration pattern was "gridline" (ie he uniformly explored the entire maritime space). Still qualitatively, even if error triangles areas greatly differed from each other, they were spread toward the periphery. As such, the global shape of the set of cognitive locations seems to be an enlarged reproduction of the shape of the actual objects. Nevertheless, some triangle superimpositions occurred.

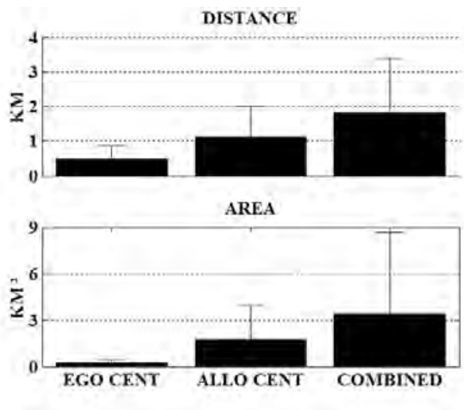


Figure 2: Means and standard deviations of the cognitive distances (KM) and the areas (KM^2) of the error triangles as a function of the frames of reference.

From quantitative point of view, the comparison of the sensory modalities showed that the cognitive distance was 1.02 ± 0.59 km when the subject explored the haptic map and 1.26 ± 1.57 km when he explored the tactile map with no significant difference ($F(1,5)=.39$; $p>.05$). In line with this result, no difference was found for the areas ($F(1,5)=.32$; $p>.05$) with 1.43 ± 2.10 km^2 for the haptic map and 2.13 ± 4.47 km^2 for the tactile one. The comparison of the reference frames is displayed on figure 2. It showed increasing values for both distance and area as a function of the reference frame. Nevertheless, whereas differences existed concerning the distance ($F(2,10)=4.52$; $p<.05$), we found no difference concerning the area ($F(2,10)=2.96$; $p>.05$). A post hoc comparison of the average distances showed that the allo centered condition do not differ ($p>.05$) from neither the ego centered nor the ego oriented one but that these two latest conditions differed form each other ($p<.05$).

4. Discussion

In the present paper, we compared the capability of blind sailor to build a situated cognitive map learned by means of either a tactile or a haptic map.
Our main results showed that similar performances could be obtained after a tactile or a haptic exploration. Even if the subject spent twice as long to explore the

haptic map than the tactile one (17 min vs. 6 min), two series of considerations may explain the reasons why he could take advantage in using the haptic map.
If we focus on the subject prior experience, we raise that subject had always explored tactile maps and could be considered as an expert. Conversely, before participating in our experiment, he never had the opportunity to use an haptic device. Despite the fact that he was a beginner in using such device, he reached a performance level equivalent to the level he obtained after years of tactile maps training. Consequently, if the subject becomes familiar with this device, we assume that he would increase his haptic perception and reduce the distance between the actual and cognitive locations.
Furthermore, we reach a limit of the technology. On the tactile map, the subject used his ten fingers during the exploration and got almost immediately a global but static perception of the layout. On the haptic map, he used the device as a single finger. Obviously, he could not manipulate ten Phantom together. Thus, with a unique device, he got a sequential and dynamic perception. This makes us suggest that during the additional time he needed, the subject memorized the pattern of movement necessary to go from an object to another.

This result reinforces the idea of the major role played by movement in the interaction between perception and action provided that the haptic interface would be natural and intuitive. Nevertheless, they also remind that further investigation is still required to disentangle the subtle influences of the frames of reference in the capability of blind sailors to build situated cognitive maps from haptic maps exploration. Finally, the subject produced more error when he answered the questions relative to the combined frame of reference. SeaTouch could be a useful software devoted to the specific training of blind subjects immersed in virtual environments that combine frames of reference.

References

[1] M.A. Espinosa, S. Ungar, E. Ochaíta, M. Blades, and C. Spencer. Comparing methods for introducing blind and visually impaired people to unfamiliar urban environments. *Journal of Environmental Psychology*, 18(3):277–287, 1998.

[2] J.F. Fletcher. Spatial representation in blind children 1: development compared to sighted children. *Journal of Visual Impairment and Blindness*, 74(10):318–385, 1980.

[3] Gaunet F. Gentaz E. L'inférence haptique d'une localisation spatiale chez les adultes et les enfants : étude de l'effet du trajet et du délai dans une tâche de complètement de triangle. *L'année psychologique*, 106:167–190, 2006.

[4] R.G. Golledge, M. Rice, and R.D. Jacobson. A Commentary on the Use of Touch for Accessing On-Screen Spatial Representations: The Process of Experiencing Haptic Maps and Graphics. *The Professional Geographer*, 57(3):339–349, 2005.

[5] D.A. Hardwick, C.W. McIntyre, and H.L. Pick Jr. The Content and Manipulation of Cognitive Maps in Children and Adults. *Monographs of the Society for Research in Child Development*, 41(3):1–55, 1976.

[6] R.D. Jacobson. Spatial cognition through tactile mapping. *Swansea Geographer*, 29:79–88, 1992.

[7] Kitchin R. Jacobson R.D. and Golledge R. Multimodal virtual reality for presenting geographic information. *Virtual reality in geography*, pages 382–400, 2002.

[8] G. Jansson and P. Pedersen. Obtaining geographical information from a virtual map with a haptic mouse. In *Proceedings of XXII International Cartographic Conference (ICC2005)*, 9-16 July, A Coruna, 2005, Spain.

[9] R.M. Kitchin and R.D. Jacobson. Techniques to Collect and Analyze the Cognitive Map Knowledge of Persons with Visual Impairment or Blindness: Issues of Validity. *Journal of Visual Impairment and Blindness*, 91(4):360–376, 1997.

[10] R.L. Klatzky. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. *Spatial cognition. An interdisciplinary approach to representing and processing spatial knowledge*, pages 107–127, 1998.

[11] J.M. Loomis, R.L. Klatzky, R.G. Golledge, J.G. Cicinelli, J.W. Pellegrino, and P.A. Fry. Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General*, 122(1):73–91, 1993.

[12] M. Rice, R.D. Jacobson, R.G. Golledge, and D. Jones. *Design Considerations for Haptic and Auditory Map Interfaces*. Cartography and Geographic Information Society, 2005.

[13] J.J. Rieser, D.A. Guth, and E.W. Hill. Sensitivity to perspective structure while walking without vision. *Perception*, 15(2):173–88, 1986.

[14] M.J. Rossano and D.H. Warren. Misaligned maps lead to predictable errors. *Perception*, 18(2):215–29, 1989.

Extending DIMPLE: a rigid body haptic simulator for interactive control of sound.

Stephen Sinclair Marcelo M. Wanderley

Input Devices and Music Interaction Laboratory, McGill University
Centre for Interdisciplinary Research in Music Media and Technology

E-mail: sinclair@music.mcgill.ca, marcelo.wanderley@mcgill.ca

Abstract

We previously introduced DIMPLE, a software environment allowing the run-time creation of a physically dynamic, haptically-enabled virtual scene using the Open Sound Control (OSC) protocol. Object properties could be requested over OSC to allow modulation of parameters for sound synthesis or visualization. However, scenes were limited to simple 3D objects and haptic interaction was limited to pushing on objects and their constraints. Here we present various recent developments to enhance DIMPLE for more intricate haptic sensations and to allow a wider variety of user interaction techniques.

1 Introduction

DIMPLE [17] is composed of a multithreaded interface between the CHAI haptic C++ framework [5] and the Open Dynamics Engine [19], with procedure calls exposed through an OSC interface for creating objects and accessing their properties.

Since the physics and haptics systems run asynchronously, the update interval of the physics simulation is independent of the haptic resolution. While haptic rendering occurs at the required 1 KHz interval necessary to present stable and stiff surfaces [12], forces accumulated on a rigid body by the haptic proxy are applied at the next physics timestep, which is configurable from 30 to 1000 Hz (usually 100 Hz).

Currently, it is possible to create simple objects such as prisms and spheres, as well as to combine these shapes into compound objects sharing a single rigid body. Constraint-based relationships between objects or between an object and the global coordinate system can be specified, allowing hinges or springs (for example) to be created.

2 Improvements

While pushing objects around and using their dynamic behaviour to modulate parameters can be interesting, we acknowledge that much of the subtlety achieved in musical interaction occurs at the sub-millimeter level, with friction, textured surfaces, and tactile vibration. For DIMPLE to be musically interesting, it is necessary to address deficiencies in these areas. We report here some recent progress towards this goal.

2.1 Vibrotactile feedback

A feature lacking previously in DIMPLE was good feedback from the audio synthesis to the haptic rendering system. It has been shown that vibrotactile feedback is important in making controllers feel "alive" and interesting to performers [11], and also that it can be critical to performance of unpredictable and non-linear instrument characteristics [4]. Thus it would be a pity to ignore this important property of virtual controllers.

Combined haptic-audio physical modelling systems such as CORDIS-ANIMA [3] make use of a single physical model, rendered synchronously, to produce both haptic rendering and audio synthesis that are tightly integrated. Haptic frequencies present in the vibration of objects at the audio rate are also felt in the haptic feedback due to this combined approach which requires executing the complete physical model in real-time at a high frequency. DIMPLE, in contrast, has taken an approach emphasizing the use of haptics for manipulation and movement, but producing audio feedback in a separated synthesis engine that is decoupled from the simulation. While this makes it simple to construct haptic virtual controllers for sound, these could not exhibit vibrations felt in response to the audio output.

The missing link has been a way of transmitting audio information back from the synthesis engine to the physical model. Audio engines such as Pd represent audio differently than other data due to its real-time nature, and thus the connection to DIMPLE should also take this into account. We previously discussed the possibility of using the JACK audio connection kit [7] to route audio data into the DIMPLE process, since it is designed to transmit audio between processes under real-time constraints. However, more recently we have ported DIMPLE to a PureData external [18], which has turned out to

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

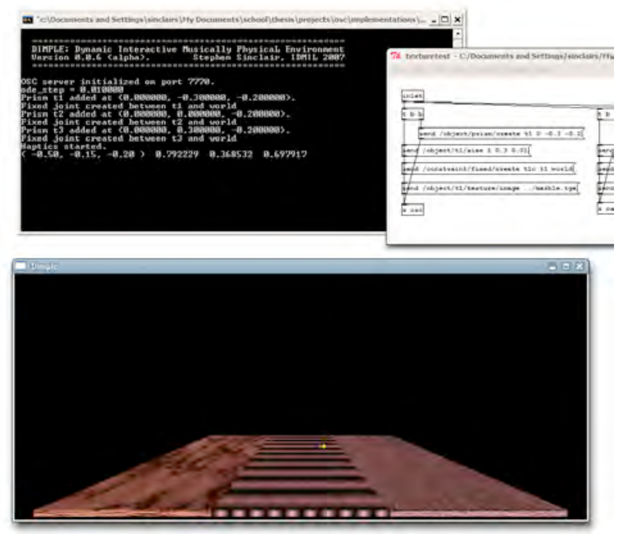


Figure 1: The PureData test patch for specifying haptic textures in DIMPLE.

be useful for this purpose. The original intention for this was to help in testing latency differences between using UDP/IP for OSC messages versus passing them between threads within shared memory. This has fortunately also made it possible to connect an audio patch cord in Pd to the DIMPLE object, giving access to the Pd DSP subsystem directly. The DIMPLE-Pd's audio handler runs samples through routines from libsamplerate [8], a resampling C library, in order to downsample the audio to the haptic rate. The audio data is then transfered through a circular buffer to the haptic thread, where it can be used to modulate the force vector.

2.2 Haptic textures

Used quite often in percussion and sound effects are interactions such as scratching, scrubbing, rolling and sliding. One way to achieve this is to perform haptic rendering of a texture, and to transmit information about the interaction stochastics to the audio synthesis engine [21].

CHAI, which is used by DIMPLE both for haptic and visual feedback, has capabilities for rendering visual textures, but lacks support for haptically representing them. We have developed an extension which modulates the stiffness of a surface based on a given height map. Though simple, this technique successfully gives an impression of textures aligned with the visual bitmap because the hand can feel oscillations correlating with visual movement.

An image of the texture interaction test can be seen in Figure 1. It shows the four messages needed to initialize each of the objects in the scene. These textures were generated synthetically using the GIMP graphics tool.

2.3 Proxy object "grabbing"

CHAI works by presenting a haptic "proxy" object, a scene element that represents the location of the force-feedback controller's end effector. The algorithm used for calculating forces on the proxy is based on the *god-object* technique [23]. Since CHAI is primarily a 3-degree-of-freedom (3-DOF) system, this proxy object is presented as a point-like sphere.

While this is adequate for touching the surfaces of objects and pushing on them, it leaves the user with a sense of merely "poking around" in the scene. In contrast, many instrumental gestures depend on interaction through a tool. Examples in music include the bow, the drumstick, or the guitar pick.

A new feature in DIMPLE is to allow "grabbing" objects. When an object is grabbed, it is no longer touched by the proxy object, but instead a stiff spring is created between the object and the proxy position. This spring acts on the object in the physics routine to pull it towards the location of the proxy, and in the haptics routine opposite forces from the spring are applied to the haptic device, pulling the end effector towards the object. This has the effect of associating the object with the proxy, such that user has a sense of directly manipulating a scene object with the haptic device. The object's inertia is felt and, if gravity is present, the object's weight is also distinctly perceptible. Collisions between the grabbed object and other objects are also felt in the haptic device.

The reason for using a spring between the haptics and physics threads instead of simply running the physical simulation at each haptic iteration is that it allows tight interaction between these two processes without sacrificing asynchrony and without requiring the demanding physics computations at haptic rates. While it may impose a lack of fidelity to the object movement, since a mass-spring system implies a certain phase delay, the response can be tuned according to the chosen physics update rate. The advantage is that the spring can be calculated at haptic rates while object movement (physics updates) can still occur at a lower speed. In some ways it is reminiscent of a 6-DOF version of the *god-object* method, since in this case the physical dynamics engine is essentially taking over the responsibility of haptic rendering.

2.4 Mesh objects

DIMPLE now makes use of CHAI's routines for loading 3D objects saved in the Alias Wavefront ".obj" and the 3D Studio Max ".3ds" formats. An equivalent triangular mesh is generated for ODE, as per the `dynamic_meshes` example given in the CHAI 1.61 download [6]. Mesh objects can be touched and interact in the physical environment just like prisms and spheres. Note however that mesh objects with large numbers of

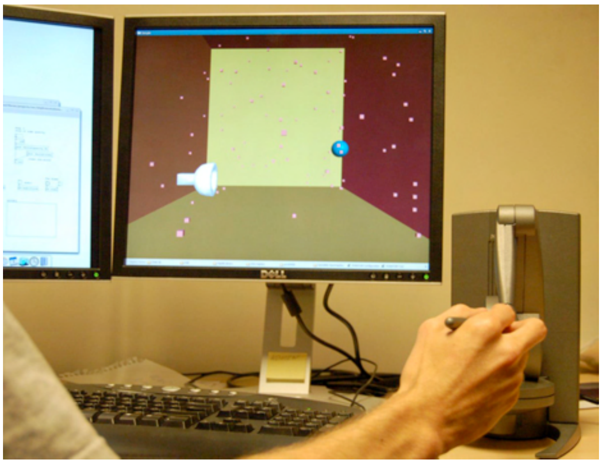


Figure 2: A demonstration of many objects bouncing around a room. The cup-shaped object is a mesh. The sphere is currently "grabbed" by the proxy, so that it is being manipulated by the haptic device. Collisions with the other objects are felt by the controller.

triangles require much more computational power than simple prisms and spheres.

Using mesh objects opens possibilities for more interesting physical interactions. Oddly shaped objects may rebound in different ways than simple spheres and cubes. Grabbing mesh objects, as described above, can provide the user with a variety of "tools" that are more interesting to manipulate than these simple objects.

3 Future Work

The improvements discussed here still require further work and verification. This section will enumerate several planned activites related to upcoming DIMPLE development.

The circular buffer used to transmit audio information from PureData to DIMPLE does not minimize haptic latency, since information must be transfered from the haptic thread to the audio thread and back. This imposes a delay of at least a few haptic timesteps (depending on the audio buffer size) before an event produces vibrational feedback. Generally, inter-modal latency (haptic events causing audio events) is more forgiving (~24 ms [1]) than purely haptic latency. Therefore, studies are required to determine whether or not DIMPLE's vibrotactile latency lies below the perceptual threshold, a requirement for providing convincing tactile simulation of virtual vibrating bodies.

The haptic texturing method employed thus far amounts to modulation of surface stiffness. This technique does not provide resistance to lateral movement as in a real texture. We hope to improve this in the future by using normal or geometry mapping, as described in [20]. We are equally interested in exploring generative textures in 2 and 3 dimensions [10], since these can be specified on the fly with a few parameters which can easily be stored in a PureData control patch, and have advantages for continuity across surfaces. Synchronizing haptic textures with shader-generated visual textures may prove an interesting challenge.

Additionally, DIMPLE does not yet send information about texture interaction to the audio process. We intend to define a set of OSC messages for sending these micro-collisions as a series of impulses, or as stochastic coefficients. Such information may also be a good candidate for using an *out-going* PureData signal connection, like the in-coming one used for vibrotactile feedback. This information could then be used for synthesis of sonic textures using one of several methods, such as that described in [2], for example.

3.1 Friction models and application to bowing

One benchmark we are using to evaluate these extensions to DIMPLE is to see if it can be used to model bowing gestures. This is because bowing has strict requirements for a satisfying simulation. Additionally, it is a useful baseline because it has been previously implemented several times [9, 13, 14], and there is therefore a good basis for comparison. In a companion paper at this conference [15], we suggest that all the components of torque play a part in bowing gestures, thus the importance of implementing basic 6-DOF support in DIMPLE.

Bowing, however, also depends on a friction model designed after the interaction between a bow hair and a string. Such friction models, based on the idea of stick-slip motion, have been previously developed [9, 16] and succesfully used in virtual haptic bowing [9, 13]. Currently CHAI supports a friction model described in [22]. This does in fact render a stick-slip-enhanced Coulomb-like friction, but it is not based on the hair-string interaction as is needed for bowing. It will thus be necessary to expand CHAI to support various other friction models.

Additionally, it is not clear how the friction haptic model might interact with the 6-DOF physics engine, since they operate independently in DIMPLE. At this time, friction simulation is lost when a "grabbed" object acts as the proxy, except for ODE's simple Coulomb friction model. It may be equally interesting to explore ways that improved friction can be incorporated into ODE.

4 Conclusion

We have described several enhancements to DIMPLE that allow more intricate interaction techniques. While the slower rigid body model in DIMPLE has advantages in terms of computational demand and allows the use of common computing hardware to maintain a real-time multi-modal simulation, it creates artificial limits on the fidelity of interaction at the texture and friction level. It is hoped that by implementing textures and friction models in the haptic thread, and by creating a pipeline between the audio synthesis engine and the haptic simulation, this fidelity can be restored.

5 Acknowledgements

This work was sponsored in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Foundation for Innovation (CFI), and the Enactive Network European Project.

References

[1] B. D. Adelstein, D. R. Begault, M. R. Anderson, and E. M. Wenzel:. Sensitivity to haptic-audio asynchrony. In *Proceedings of the 5th International Conference on Multimodal Interfaces*, pages 73–76, Vancouver, BC, November 2003. ACM.

[2] S. Barrass and M. Adcock. Interactive granular synthesis of haptic contact sounds. In *Proceedings of the AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio*, pages 270–277, Helsinki University of Technology, Espoo, Finland, June 2002.

[3] C. Cadoz, A. Luciani, and J. L. Florens. CORDIS-ANIMA: a modeling and simulation system for sound and image synthesis—the general formalism. *Computer Music Journal*, 17(1):19–29, 1993. MIT Press.

[4] C. Chafe. Tactile audio feedback. In *Proceedings of the International Computer Music Conference*, pages 76–79. ICMA, 1993.

[5] F. Conti, F. Barbagli, R. Balaniuk, M. Halg, C. Lu, D. M. L. Sentis, E. Vileshin, J. Warren, O. Khatib, and K. Salisbury. The CHAI Libraries. In *Eurohaptics '03*, pages 496–500, Dublin, Ireland, June 2003.

[6] F. Conti, D. Morris, F. Barbagli, and C. Sewell. CHAI 3D. Available: `http://www.chai3d.org/`, November 2006.

[7] Davis, P. et al. JACK Audio Connection Kit (software). Available: `http://jackaudio.org`.

[8] E. de Castro. libsamplerate (software), August 2007. `http://www.mega-nerd.com/SRC/`.

[9] J.-L. Florens. Expressive bowing on a virtual string instrument. In A. Camurri and G. Volpe, editors, *Lecture Notes in Computer Science: Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop*, volume 2915, pages 487–496. Springer Berlin / Heidelberg, Genova, Italy, April 2003.

[10] J. P. Fritz and K. E. Barner. Stochastic models for haptic texture. In M. R. Stein, editor, *Proceedings of SPIE*, volume 2901, pages 34–44. SPIE, 1996.

[11] B. Gillespie. Haptics. In P. R. Cook, editor, *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*, chapter 13, pages 229–246. MIT Press, Cambridge, MA, March 1999.

[12] M. Minsky, O.-Y. Ming, O. Steele, F. P. Brooks Jr., and M. Behensky. Feeling and seeing: issues in force display. In *SI3D '90: Proceedings of the Symposium on Interactive 3D Graphics*, pages 235–241, New York, NY, USA, 1990. ACM Press.

[13] C. Nichols. The vBow: development of a virtual violin bow haptic human-computer interface. In *Proceedings of the Conference on New Interfaces for Musical Expression*, pages 29–32, Dublin, Ireland, 2002.

[14] S. M. O'Modhrain. *Playing by Feel: Incorporating Haptic Feedback into Computer-Based musical Instruments*. PhD thesis, Stanford University, November 2000.

[15] E. Schoonderwaldt, S. Sinclair, and M. M. Wanderley. Why do we need 5-dof force feedback? the case of violin bowing. In *Proceedings of Enactive '07*, 2007.

[16] S. Serafin, C. Vergez, and X. Rodet. Friction and application to real-time physical modeling of a violin. In *Proceedings of the International Computer Music Conference*, pages 375–377, Beijing, China, 1999. ICMA.

[17] S. Sinclair and M. M. Wanderley. Defining a control standard for easily integrating haptic virtual environments with existing audio/visual systems. In *Proceedings of the Conference on New Interfaces for Musical Expression*, pages 209–212, New York, NY, June 2007.

[18] S. Sinclair and M. M. Wanderley. Using PureData to control a haptically-enabled virtual environment. In *PureData Convention '07*, Montreal, Canada, August 2007.

[19] R. Smith. Open Dynamics Engine (software). Available: `http://www.ode.org`, November 2006.

[20] V. Theoktisto, M. Fairén, I. Navazo, and E. Monclús. Rendering detailed haptic textures. In F. Ganovelli and C. Mendoza, editors, *Workshop On Virtual Reality Interaction and Physical Simulation*, pages 8–16. Eurographics Association, 2005.

[21] K. van den Doel, P. G. Kry, and D. K. Pai. FoleyAutomatic: physically-based sound effects for interactive simulation and animation. In *Proceedings of the 28th annual conference on computer graphics and interactive techniques*, pages 537–544, 2001.

[22] C. Zilles. Haptic rendering with the toolhandle haptic interface. *Master's thesis*, Massachusetts Institute of Technology, Cambridge, MA, 1995.

[23] C. Zilles and J. Salisbury. A constraint-based god-object method for haptic display. In *Proceedings of the International Conference on Intelligent Robots and Systems*, volume 3, page 3146, 1995.

Dynamic anamorphosis

Franc Solina   Borut Batagelj

University of Ljubljana, Faculty of Computer and Information Science,
Computer Vision Laboratory, Tržaška 25, 1000 Ljubljana, Slovenia
E-mail: {franc.solina, borut.batagelj}@fri.uni-lj.si

Abstract

In this paper we define the concept of dynamic anamorphosis. A classical or static anamorphic image requires a specific, usually a highly oblique view direction, from which the observer can see the anamorphosis in its correct form. This paper introduces dynamic anamorphosis which adapts itself to the changing position of the observer so that wherever the observer moves, he sees the same undeformed image. This dynamic changing of the anamorphic deformation in concert with the movement of the observer requires from the system to track the 3D position of the observer's head and the recomputation of the anamorphic deformation in real time. This is achieved using computer vision methods which consist of face detection/tracking of the selected observer and stereo reconstruction of its 3D position while the anamorphic deformation is modeled as a planar homography. Dynamic anamorphosis can be used in the context of art installation, in video conferencing to fix the problem of the missing eye contact and can enable an undistorted view in restricted situations.

1. Introduction

In this paper we define the principle of dynamic anamorphosis which adapts itself to the changing position of the observer in such a way that he always sees the image in its correct un-deformed form. First, we introduce the classical principle of anamorphosis. Next, we extend this principle to dynamic anamorphosis. Later sections describe in some more detail the required technical background for implementing a system which uses dynamic anamorphosis. Computer vision methods are used to determine the 3D position of the observer's head. The anamorphic deformation of the displayed image is computed as a planar homography. In Conclusions we discuss possible user scenarios for dynamic anamorphosis.

2. Anamorphosis

Anamorphosis or anamorphic projection was discovered in art the late fifteenth century both as a challenge and as a confirmation of the rules of linear perspective which were discovered at the same time [5]. Classical linear perspective is based upon the Euclidean paradigm that light travels in straight lines and when light reflected from an object intersects a planar surface an accurate representation of the original object is reflected on that surface. While we normally look at images frontally from a limited range of viewing angles, the viewer of an anamorphic image must usually be at a radically oblique angle to the picture plane to see the anamorphic image undistorted. The anamorphic image looked at up front is in such cases usually so distorted as to be unrecognizable.

Figure 1: The Ambassadors by Hans Holbein, 1533, Oil on oak, 207 × 209 cm, National Gallery, London.

Probably the most famous example of anamorphosis in art history is the 1533 painting *The Ambassadors*, by Hans Holbein (Fig. 1). On the bottom of this painting appears a diagonal blur which appears as a human skull when viewed from the upper right [12]. Therefore, anamorphosis was traditionally used to depict subjects

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

one was reluctant to represent directly such as erotic, occult or otherwise controversial imagery. Probably the first known anamorphosis is from Leonardo de Vinci's Codex Atlanticus (1483–1518) and shows the face of a child [12].

Somewhat later, in the sixteenth century, other types of anamorphosis were developed, such as the so-called cylindrical anamorphosis which requires a cylindrical mirror to observe the image correctly. Famous, for example, is a cylindrical anamorphosis of Jule Verne's portrait drawn by the Hungarian artist István Orosz. The image looks up front as a shipwreck in an arctic landscape. By placing a cylindrical mirror on the sun depicted in the image, Jule Verne's face suddenly appears. Due to the ease of producing anamorphic images using computer graphics they appear now often in newspapers and magazines. Anamorphic images are even produced as pavement art [2]. Special examples of anamorphic projection are also images which are projected on slanted surfaces (pavements, ceilings) but appear undeformed.

Since the appreciation of anamorphic images requires an "eccentric" viewing point as opposed to a "normal" or orthogonal viewing point, anamorphosis is a term popular with many postmodern theorists used mainly as a metaphor for the relativity of vision or the subjectivity of human experience [12]. Anamorphosis serves as a model for the concept of the gaze, which suggests that visual appreciation rather then passive "looking" requires active "observing" [8, 14].

To appreciate an anamorphic image requires indeed from the observer that he positions himself precisely in the right spot and directs his gaze in the right direction as opposed from the "normal" or "centric" vision [1] where the viewer sees himself at the center of the world and as he moves, the center of the world moves with him and the world surrounding him stays coherent. To view an anamorphic image one has to transform an oblique and non-uniform focal plane into a coherent, two-dimensional image which is sometimes facilitated by viewing with one eye or with half-closed eyes [5]. This enables the dissociation of the image from the screen or the supporting surface and the anamorphosis re-forms itself [12]. Viewing "normal" pictures from an oblique angle does not result in a distorted picture since human perception can automatically compensate for the distortion using the principle of shape constancy. Straying away from the right viewpoint of an anamorphic image, on the other hand, can quickly deteriorate the effect.

3 How dynamic anamorphosis works

To see a static anamorphic image one has to position oneself in the right spot and then view the image in the right direction. Since we plan to project the anamorphic image using a video projector which is connected to a computer, we can reshape the anamorphic image whenever the observer moves in such a way that the re-formed image stays for the observer the same. To achieve this constancy of the re-formed anamorphic image one has to track the position of the observer in real-time and then according to the established position pre-deform the projected anamorphic image in real time so that it appears un-deformed from that particular view point. If traditional anamorphosis requires an accurate, often "eccentric" viewpoint, this installation uses anamorphosis to separate the human spatial orientation from the visual cues and can thus provoke a crisis in the visual faculty—wherever the observer moves in space, he sees the same re-formed image.

A somewhat similar concept involving imagery that adapts to the position of the viewer is described by Steve Mann [10]. The observer wears special eyeglasses that track where the person is and then the system generates stabilized images on displays to sustain the illusion of a transparent window showing the subject matter behind the display in exact image registration with one would see it if the display were not present.

3.1 Localization of the observer

To drive the anamorphic projection we need to know the position of the observer. The most unobtrusive technology to determine the position of objects in a given scene is provided by computer vision. We use a face detection method to determine the position of the user's face in the pictorial plane. By using two or even more cameras and the principle of stereo reconstruction of distances we can further determine the position of the user's head in 3D space. Face detection is now a mature technology and methods such as the one developed by Viola and Jones [13] can run in real-time. The most difficult problem in stereo reconstruction is the correspondence problem—to find for a given point in the left image the corresponding point in the right image [7]. Since the number of possible matches goes into thousands of points this is a computationally intensive task. The correspondence problem in this particular case will be solved by finding faces in both images first. Next, only correspondences between faces needs to be established.

Face detection can be performed based on several cues: skin color, motion, facial/head shape, facial appearance, or a combination of these parameters. Most successful face detection algorithms are appearance-based. The processing is done as follows: An input image is scanned at all possible locations and scales by a subwindow. Face detection is posed as classifying the pattern in the subwindow either as a face or a nonface. The face/nonface classifier is learned from face and nonface training examples using statistical learning methods.

For our purpose we used the AdaBoost [13] learning-based method because it is so far the most successful in term of detection accuracy and speed. AdaBoost is used to solve the following three fundamental problems: (1) learning effective features from a large feature set; (2) constructing weak classifiers, each of which is based on one of the selected feature set; and (3) boosting the weak classifiers to construct a string classifier. Weak classifiers are based on simple scalar Haar wavelet-like features, which are steerable filters. We use the integral image method for effective computation of a large number of such features under varying scale and location, which is important for real-time performance. Moreover, the simple-to-complex cascade of classifiers makes the computation even more efficient, which follows the principles of pattern rejection and coarse-to-fine search. In the case if the cameras are placed laterally we use an extended Haar feature set for dealing with out-of-plane (left-right) rotation.

With the proposed face detection method we get the location in the image plane of all the present faces regardless of their position, scale, orientation and age. To improve the detection in the case of low illumination we also use near-infrared cameras, which are invariant to illumination and capture the temperature of the body. For head tracking we must detect the faces in every frame. To improve the tracking we use addition clues such are motion, skin color or near-infrared image.

To locate the 3D position of the detected faces in the scene we use the stereo paradigm. Most stereo reconstruction methods are based on a punctual correspondence, but they are inefficient in most realistic contexts [11]. We approach the stereo matching problem as a matching between homologous faces, instead of point matching. The main idea is to determinate a unique disparity value for the whole face region and no longer for individual pixels. After we detect the position of faces in both stereo images using the above described face detection method we construct a graph for each image where face blobs are represented as nodes in the graph. To find homologous faces in both stereo images we perform graph matching. A bipartite graph matching between the two graphs is computed in order to find for each face blob in the left image the corresponding face blob in the right image. Using the horizontal displacement between the corresponding face blobs, the position and other calibration parameters of the stereo cameras, we can computer the 3D position of each face in the scene.

The computational process is simple and fast since we consider only complete face regions. The result is robust in a realistic context, because an integral measurement of the disparity for the whole face region can mitigate some local and global fluctuations between the stereo image pair.

3.2 Anamorphic deformation

Recent computer-controlled video projection systems have one or more built-in cameras to provide a visual feedback that can automatically compensate for the so called "keystone" deformation. The keystone deformation can be represented in the most general way as a planar homography mapping points in the projector plane onto the screen plane, corresponding to a 3-degrees of freedom alignment (pan, tilt, screw) [3]. To eliminate the effect of the keystone, its associated homography can be estimated and used to suitably pre-deform the image being displayed.

The same homography can be used to make a "virtual anamorphosis" so that the image is intelligible only for observers looking at the screen from a particular viewpoint [3]. The authors call this functionality directional vision and compare it to directional audio. We use this homography to deform the projected image of the face in such a way that it looks undeformed from the viewpoint of the observer.

4 Applications of dynamic anamorphosis

The first application of dynamic anamorphosis was in the context of an art installation where the subject of the projected image is a human face looking straight ahead. People are very sensitive to the direction of the eye gaze. Our eyes express our emotions and intentions and they help us direct attention. Cultural norms dictate when, for how long and in what situations it is appropriate to gaze into another persons eyes. We can determine very accurately if somebody is actually looking at us. Eye gaze is important in conversations and the lack of proper eye contact in videoconferencing systems is a serious limitation of such systems [4]. When a participant in a videoconference watches the images of other participants on the video monitor he cannot look into the camera at the same time even if the camera is placed just above the monitor.

In the installation the face with the eye gaze turned directly ahead will meet the eyes of the installation user. Due to the viewpoint-sensitive anamorphic deformation the projected face in the installation will stare at the installation user wherever he moves. There will be no way to escape it's gaze. This should be for most installation users a rather unnerving situation. On a symbolic level the installation epitomizes the personification of ubiquitous video Surveillance systems [9]. Instead of a static image a short video loop of a face gazing straight ahead can be used such as the Big Brother from the film after George Orwell's novel 1984. Figure 2 shows the transformed frame of video clip, when the user views it under 30° angle from right.

The installation requires a dark room with the video projection over an entire wall so that the only visible

Figure 2: The transfomed image of the Big Brother from the film after George Orwell's novel 1984.

cues seen by the user are given by the projection. The video clip used for projection should feature a human face with the eye gaze directed straight ahead. The light reflected back into the room from the projected image must sufficiently illuminate the scene that face detection can be performed.

Since the installation can truly be experienced only by a single user the entrance to the room with the installation should be controlled. Another possible scenario for the exhibition would allow several people in the audience. Of course, only one of them must somehow be selected for the virtual anamorphic experience using either face recognition or some other distinguishing feature used for stereo matching. Other people in the room could enjoy the projection which will be deformed from their viewpoint or they will try to move into the position from which the anamorphosis will re-form itself.

Other uses of dynamic anamorphosis can be envisioned related either to fixing the problem of the missing eye contact in video conferencing or how to enable an undistored view of visually conveyed information in restricted situations.

5 Conclusions

We introduced the concept of dynamic anamorphosis which enables an undistorted view of an image from almost any position. Classical or static anamorphosis forces the observer to find the special viewpoint from which the anamorphic image is re-formed. Dynamic anamorphosis instead disassociates the geometric space in which the user moves from the visual cues he sees, since wherever the observer moves, he sees the same image.

The principle of dynamic anamorphosis was initially developed for an art installation where a human face with the eye gaze directed straight ahead to meet the eyes of the installation user was selected for the projected image. For the real-time tracking of the observer's face we use computer vision methods. The 3D position of the face in turn also determines in real-time the homography that deforms the anamorphic projection in such a way that the projected face that the observer sees is constantly gazing towards him.

References

[1] R. Arnheim. *The Power of the Center*, University of California Press, Berkeley, 1988.

[2] Julian Beever. Pavement drawings, http://users.skynet.be/J.Beever/pave.htm, May 2006.

[3] A. Brazzini, C. Colombo. Computer Vision for Interactive Skewed Video Projection. In F. Roli, S. Vitulano (Eds.): *Image Analysis and Processing - ICIAP 2005*, 13th International Conference, Cagliari, Italy, September 6-8, 2005, Proceedings. Lecture Notes in Computer Science 3617, Springer 2005, pp. 139-146.

[4] R. A. Colburn, M. F. Cohen, S. M. Drucker. The Role of Eye Gaze in Avatar Mediated Conversational Interfaces. Technical Report MSR-TR-2000-81, Microsoft Research, Microsoft Corporation, 7/31/2000.

[5] D. L. Collins. Anamorphosis and the Eccentric Observer: Inverted Perspective and Construction of the Gaze. *Leonardo* 25(1): 72–82, 1992.

[6] D. L. Collins. Anamorphosis and the Eccentric Observer: History, Technique, and Current Practice. *Leonardo* 25(3): 179–187, 1992.

[7] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, 1993.

[8] J. Lacan. *Four Fundamental Concepts of Psycho-Analysis*. W. W. Norton, New York, 1978.

[9] T. Y. Levin, U. Frohne, P. Weibel. *CTRL [SPACE], Rhetorics of Surveillance from Bentham to Big Brother*. (Karlsruhe: ZKM and Cambridge: MIT Press, 2002).

[10] S. Mann. *Intelligent Image Processing.* Wiley-IEEE Press, 2001, Chapter 3.

[11] D. Scharstein, R. Szeliski. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. *International Journal of Computer Vision* 47(1-3): 7–42, 2002.

[12] D. Topper. On Anamorphosis: Setting Some Things Straight. *Leonardo* 33(2):115–124, 2000.

[13] P. Viola, M. J. Jones. Robust Real-Time Face Detection. *International Journal of Computer Vision* 57(2): 137–154, 2004.

[14] S. Žižek. *The Sublime Object of Ideology*. Verso, London, New York, 1997.

Is "Life" Computable?
On the Simulation of Closure under Efficient Causation

John Stewart[*] Matteo Mossio[*†]
(*) Groupe Suppléance Perceptive, Université de Technologie de Compiègne, France
(†) REHSEIS, Université Paris 7, France*
E-mail: john.stewart@utc.fr, matteo.mossio@ens.fr

Abstract

In this paper, we discuss Robert Rosen's conjecture according to which a system closed under efficient causation possesses noncomputable models. We suggest that this conjecture relies on a fundamental misunderstanding between circularity and noncomputability, and it is therefore false. Nevertheless, Rosen's work contains several powerful insights about the theoretical distinction between simulation and modeling. We discuss some major implications of this dichotomy for the community of Artificial Life and the enactive approach to Life and Cognition.

1. Introduction

We will begin by three preliminary remarks.

The first remark is that living organism itself does not compute, of course, any more than the planets compute their velocities as they go round the sun. The proper question is whether living organisms have Turing-computable models. This means that we will have to look carefully at what is involved in a modeling relationship between a natural system and a mathematical model; and in particular, to distinguish between modeling causality on one hand and mere imitation by computational simulation on the other; a theme that is dear to Giuseppe Longo [7].

The second remark is that if we merely adopt a common-sense definition of life, it is obviously possible to build computational models of certain aspects of living organisms. For example, it is possible to model the flow of blood in arteries and veins, using the same hydrodynamics as for the flow of water in a tube. However if we do this, we are in an important sense missing the point: we are not modeling the organism as living, we are just treating it as thought it were not alive [13]. Thus, to answer our question, we will have to provide a theoretical model which captures the essence of "life itself", to use Rosen's expression. And the question will then be whether this model is computable.

The third remark is that the question has not only a great theoretical interest, but also it is of considerable practical importance, because of the existence of a substantial scientific community known as "Artificial Life". In this community, more than 90% of the work in this field consists of computer simulations. If it were to be shown that a model of "life itself" were not computable – what we shall henceforth call "Rosen's conjecture" – this would not disqualify the field, but it would have major impact. We may note here that if the question is properly posed, it would seem intuitively at least that it must have a clear-cut answer, either "yes" or "no". The first aim of this paper is quite modestly to pose the question as precisely and correctly as possible; and to indicate some of the current lines of research aimed at providing an answer. The second aim, more ambitious, is to set forth an answer outline, so that it may be subjected to critical examination by the scientific community.

2. Fundamental Theories of Life

The heart of the question lies, as we have already indicated, in a theoretical definition of life. In contemporary biology, the central scientific object, which is properly constituted in theory, is not "life" but the gene [4]. The consequence is that contemporary molecular biologists declare roundly that "life does not exist"; or, more prudently, that "life is not an object of study in biological laboratories" [15].

At the present time, there are two major proposals aimed at providing a fundamental theory of life: the concept of autopoïesis [10], and the concept of closure under efficient causation [14]. Although these theories were developed quite independently of each other, they have important features in common: notably, metabolism and circularity at an abstract level of organization, such that the set of processes occurring in the system collectively ensures the conditions for their own continuation. A certain amount of current research is aimed at clarifying the relation between these two theories ([5], [12], [17]). The general consensus so far is that these theories substantially overlap; to the extent that they are different, they are not incompatible but rather complementary. There is also discussion as to whether autopoïesis and/or closure under efficient cause are fully sufficient for defining life [1]; but there is a reasonable consensus that they are at least necessary, so that if they have non-computable models then life will too.

In this paper, we shall focus on Rosen's formulation in terms of "closure under efficient causation", for several reasons. Firstly, although the concept of autopoiesis is intuitively easier to understand, and has thus had a wider influence not only in biology but beyond in the field of social theory, it has not (yet?) received an adequate mathematical formulation. Secondly, Varela himself has proposed a "tesselation automaton", eminently realizable by computer simulation, as an illustration of the concept of autopoïesis. To put it mildly, this does not suggest that there is anything problematical about a computational model of autopoïesis; so this formulation is not particularly appropriate for our present purpose. By contrast, Rosen does propose a mathematical formulation, in terms of Category Theory; and he claims quite explicitly that any material realization of a system exhibiting closure under efficient cause *must have noncomputable models*. It is true that Rosen's book is not easy to understand; in addition, it is riddled with many rather trivial but annoying errors and suffers from an astounding absence of concrete illustrations [6]. However, these are defects which are eminently open to correction, and many authors are making considerable contributions in this direction.

3. Closure under efficient cause

Rosen's formulation is based on a rehabilitation and reinterpretation of the Aristotelian categories of causality: material cause, efficient cause and (but we will not discuss this point here), under certain conditions, final cause. Given a mathematical function, b = f (a), there are two answers to the question "why b?": (i) "because a", i.e. the argument of the function, which Rosen interprets as the "material cause"; and (ii) "because f" where the function f is interpreted as the "efficient cause". In set-theoretical terms, a and b are sets, and f is the mapping from a to b. Applied to the case of state-determined dynamic systems (SDDS), the mapping is an endomorphism from x (t), the state-vector at time t, to the "next state" x (t+δt). The whole art of finding a mathematical expression of SDDS is to choose the state variables in such a way that the state x (t+δt) is a function only of the state x (t).

This formulation enables Rosen to express what is, in his view, the difference between physics and biology. In physics, we can ask questions about the state of a dynamic system. "Why x (t)?" – (i) "because x(t0)", the state at any reference time t0; this is the "material cause"; and (ii) "because f", the dynamic law; this is the "efficient cause". But if we ask the question "why f?", within physics there is not really any answer, other than that this just is a natural law. Rosen's proposition is that this is where biology is different: for a living organism, the question "why f?" has a non-trivial answer from within the functioning of the system itself. Let us look at this a little closer.

As we have said, there is fairly wide agreement that metabolism is at the core of living organisms. In Rosen's formalism, this is expressed by the formula:

B = f(A) Equation [1]

where A is the "material cause" of the metabolism, f is the "efficient cause" of the metabolism, and B is the result. To give a rough-and-ready interpretation, A corresponds to the input materials and energy; f may be associated with the set of enzymes which are necessary to catalyse the biochemical reactions, but also the cell membrane, necessary to avoid loss of reactants by diffusion, and probably other features of cell organization as well; and B corresponds to the total resulting biochemical network. However, as we have said, concrete illustrations are not Rosen's strong point.

What characterizes living organisms is that the maintenance, and indeed the ongoing production of this "metabolism" function, are themselves ensured by the functioning of the organism. In Rosen's formalism, this is expressed by a second function, Φ, which takes suitable input X as its "material cause" and which produces f:

f = Φ (X) Equation [2]

Now by an interaction of the same argument, we must now ask: "why Φ?" As before, we can introduce a new function, β, which Rosen calls "replication":

Φ= β (Y) Equation [3']

The point is that we now see clearly an incipient infinite regress; on the face of it we will require another function for the production of β, and then yet another function for the production of this function, and so on indefinitely. We come now to the key point: Rosen makes the crucial postulate that the infinite regress can be avoided by introducing a circularity: β is none other than B, which is already produced by the system (Equation [1]):

Φ= B (Y) Equation [3]

We thus arrive at the situation which Rosen calls « closure under efficient causation », as illustrated in Figure 1. The three efficient causes – f, the metabolism function; Φ, the repair function; and B, the replication function – are all produced by the operation of the system itself. Rosen considers that closure under efficient causation is the essential defining characteristic of “life itself”.

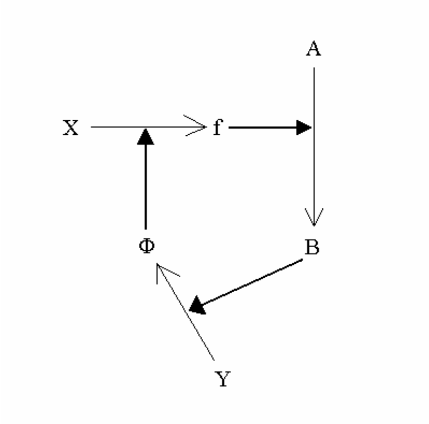


Figure 1. A relational model of closure under efficient cause. Thin lines with open arrows represent relations of material causation ; thick lines with solid arrows represent relations of efficient causation.

4. Modeling and simulation

We come now to the question we seek to pose in this paper: can the relational model illustrated in Figure 1 be implemented by a computer simulation? In order to address this question effectively, we need to be more precise as to what we mean by “modeling”, and what we mean by “simulation”.

Rosen [14] defines the modeling relationship by the diagram in Figure 2. If the diagram commutes, so that (2) + (3) + (4) always gives the same result as (1), then causal entailment in N has been brought into congruence with inferential entailment in F. In this case, since F is a human construction and therefore a priori intelligible, scientists will have understood something about the natural system N.

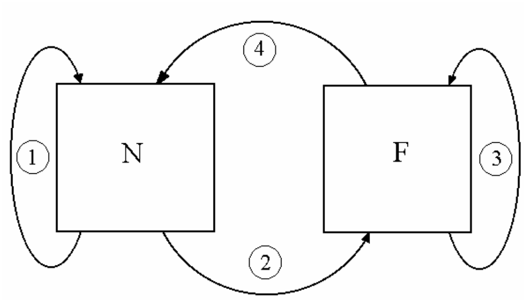


Figure 2. The modeling relationship between a natural system N on the left, and a formal system F on the right. The arrow (1) corresponds to causal entailment in N; the arrow (3) corresponds to inferential entailment in F. The encoding arrow (2) from N to F corresponds to measurement; the decoding arrow (4) from F to N corresponds to predictions about N. After Rosen [14, p.60].

Rosen bases his definition of simulation on a relational model of (any formal equivalent of) a Turing machine. If such a mathematical machine (i.e. a digital computer) is considered itself as a natural system, one can make a principled and watertight distinction between the hardware (i.e. the reading-head, and more generally the actual physical implementation of the computer, nowadays made of electronic components), and the software (i.e. the tape and the symbols inscribed thereon). In Rosen’s terminology, the hardware corresponds to the efficient cause of processes occurring in the machine, and the software to the material cause. Given a set of symbols or an alphabet A, and the set A# of words (or symbol strings) on that alphabet, the operation of a Turing machine can be considered as a mapping

f: A# -> A#

defined by

f(w) = w*

where w is the input data and w* is the output. We repeat that w is the material cause, and f the efficient cause of the operation.

Now, let us take a word u in A#. We can then define a new mapping, gu : A# -> A#, by writing

gu(w) = f(wu)

for all words w in A# .

Rosen [14, p. 192] writes:

“We can trick the mapping f, and the algorithm that computes its values, into evaluating a different map (gu). This trick is of the essence of simulation and program. In general, we shall say that f simulates g if there is a word [u] in A# such that g = gu [for all w]. …. Let us call the word u the program for the mapping gu. In a certain sense, the program u is a description of gu, recognized by the mapping f.”

Now the theory of Turing machines states that the map gu actually corresponds to a different Turing machine, with its own reading-head, i.e. a different efficient cause. Rosen [10, p. 193] continues:

“A Turing machine’s reading-head constitutes hardware…., the inferential machinery that entails words from other words. The words themselves, on which this machinery operates, are software. In particular, the program u is software. The simulation of a map g by a map f thus requires a… description of g as software, as input”.

It is this fact which gives the simulation relation its unique features. The hardware of g, the entailment structure that evaluates g, is not put into correspondence with the hardware of f, but rather is imaged in its software; in the input to f. Heretofore, it is precisely the entailment structures that are put into correspondence; that in fact is the very essence of the modeling relation. But in simulation, the hardware of the simulator needs to have nothing whatever to do with the hardware of what it simulates.

In causal terms, simulation involves the conversion of efficient cause, the hardware of that being simulated, into material cause in the simulator. In essence, this means that one can learn nothing about entailment by looking at a simulation. It also implies that the "time variable", which counts the number of applications of an algorithm to initial data before it halts, is completely unrelated to any corresponding "time variable" in what is being simulated.

This characterization of simulation starkly emphasizes the limits of computer simulation. In a Turing machine (or any formal equivalent thereof), there is one and only one efficient cause: the reading-head. Since there is not (and foreseeably never will be) a human-built Turing machine such that its operation actually produces its own reading-head, nothing in the operation of a computer can be adequate to authentically model closure under efficient cause. In this sense, we can already give a clear and resounding answer to our question: no, *life is definitely not computable*.

However, this is not quite the end of the story. We should indeed be wary of discarding computer simulation altogether; it is such a powerful and cost-effective tool for scientific research, and it is not for nothing that the bulk of research in the field of "Artificial Life" depends on it. Certainly, computer simulation cannot be used to *model* the closure under efficient cause of Figure 1; but might it not be possible, nevertheless, to *simulate* Figure 1 in a computer? We shall examine this question in the next two sections.

5. Simulating the production of an efficient cause

Before examining the key question as to whether Figure 1 could be simulated in a computer, we must first put in place the simulation of a more elementary situation, the production of just one or two efficient causes without closure. If we could not do this, then the problem would not lie in closure as such, but merely in this incapacity to simulate the production of an efficient cause. If we can do it, on the other hand, we will then be in a situation to examine the possibility of extending the simulation to incorporate closure. In this section, we will thus examine if we can simulate the production of f, i.e. f = Φ (X) as in Equation [2]. To be sure that we are treating f as an efficient cause, we will couple this to a simulation of Equation [1], B = f (A). But, in this section, we will deliberately avoid extending to Equation [3] which is what gives rise to closure under efficient cause.

As we have said in the previous section, in a simulation relation between f and g, the efficient causation at work in g is not put into correspondence with the efficient cause of f; rather, the efficient cause of g is imaged in the software, in the input to f. Thus in fact, if we are attempting to simulate all or part of Figure 1, genuine "efficient causes" (the reading-head of the Turing machine) will not come into the picture at all. Rather, the "efficient causes" (for the moment, just Φ and f) will appear in the guise of their software images – i.e. in their programs. In a strict sense, all the software has the status of material causation. Nevertheless, it will be essential to maintain the spirit of the distinction between "efficient cause" and "material cause"; and to do this, we will have to establish a distinction between two sorts of software, the program (the stand-in for efficient cause), and the data on which it operates.

This question has a long and fascinating history in computer science. The realization that programs and data are both software, in the last resort nothing other than a series of 0's and 1's, so that one can overcome a rigid, watertight distinction between them, was one of the great moments in this history; it was at the origin of the von Neumann architecture, which brought about a revolution in programming languages and is still employed in practically all computers today. And as we shall see, it will be essential for our purposes that what in one context belongs to the realm of data can, in another context, be employed as a program. But although the distinction can be blurred, there is no difficulty, if that is what we want, in setting up a clear-cut conceptual distinction between program and data. Indeed, before von Neumann's innovation – right from the very first computer, the Babbage machine, up until the first electronic computers in the 1940's, and including of course the purely conceptual Turing machine – program and data were not only conceptually distinct, but actually physically instanciated in characteristically different ways.

How then can we instanciate a simulation of the pair f = Φ (X) (Equation [2]) and B = f (A) (Equation [1])? The answer is not very difficult. The program Φ produces output data f of the basis of input data X. Then, f is employed as a program, taking input data A to produce output B. The key insight of this paper is our realization that this is not a mere abstract conception; it is just what computer software engineers spend their professional lives actually doing. It is well known that computers have a clean hierarchical architecture with clearly defined levels. At the bottom, there are the actual physical components – electronic transistors, silicon chips, weak electric currents and so on. On this are built discrete, symbolic 0's and 1's and logical gates. Then come operating systems, and machine languages; then programming languages such as Fortran, Basic, Pascal, C, Java…. Finally come the software programs that are what most of us actually use – Word, Excel and so on. Passing from one level to another – a task that is accomplished by the compiler –

corresponds quite precisely to using one program Φ to produce output data f, and then using f as a program, taking input data to produce output B; and so on, over what is now quite a large (but finite) number of steps.

Thus, there is no problem at all in setting up a chain of programs (simulated surrogates for efficient causes) which produce each other.

6. Simulating closure?

We are now in a position to rephrase Rosen's conjecture in precise terms. Is it possible for a chain consisting of a program which produces a program which produces a program…. (this is the simulation version of a chain of efficient causes) at some point to fold back on itself to form a *closed* organization?

Rosen's conjecture is that the answer is "no". His main argument (clarified and developed by other authors) is that closure under efficient causation is an irredeemably hierarchical cycle which, by manifesting impredicativities, cannot be replaced with finite syntactic algorithms, and then they are not computable [8]. The logic underlying the argument is that, whereas it is quite possible to use the output from a low-level program to constitute the program of the (adjacent) higher-level process, and indeed this is what happens all the time during compilation; but it is not possible to use the output from a high-level program to constitute (or even to repair) the program of a lower-level process, because the latter presupposes the former.

However, we submit that this argument is false. The reason is that it relies on a fundamental misunderstanding according to which circularities would generate noncomputability. In fact, it is quite the opposite. From the fact that, in practice, computers possess a hierarchical structure (often for security and stability reasons) it does not follow that hierarchical closed cycle cannot be computed in principle. As we will show in our talk, Lambda Calculus, the most powerful formalism for calculability, is perfectly able to describe recursive functions which can be used to adequately generate Rosen's closed diagram. In other words, circularity and impredicativity are at the very heart of computability, to the extent that recursion itself is a form of circularity. Consequently, closure under efficient causation does not represent a theoretical problem for calculability.

7. Conclusions and perspectives

In this paper, we provided the outline of an argument claiming that what we have called "Rosen's conjecture" – i.e. that a model of life defined as closure under efficient cause cannot be instanciated by a computer simulation – is false. Despite the recent renewal of the debate about Rosen's conjecture [2,3,9,16], we submit that the nucleus of the whole discussion is flawed, since closure and impredicativity are perfectly compatible with computation.

This conclusion is specifically relevant for the research field of Artificial Life which, we confidently anticipate, will continue to flourish by attempting, for example, to instantiate computational autopoiesis [11]. Nevertheless, given Rosen's distinction between simulations and models (which , we believe, maintains all its force and pertinence), one may wonder what kind of information about *natural* living systems can be provided by computational simulations, which can just "image" causal regimes at work in living beings. From the point of view of a full-scale enactive approach, the fundamental challenge consists in obtaining an adequate *model* (which still lacks) of the specific organization underlying life and cognition.

Computations simulations of closed causation, even if they are possible, could be largely irrelevant with respect to this issue.

8. References

[1] Bourgine, P. & Stewart J. "Autopoiesis and Cognition", Artificial Life, 10, 327–345, 2004.

[2] Chu, D., & Ho, W.K., "The localization hypothesis and machines", Artificial Life, 13, 3, 299-302, 2007.

[3] Chu, D., & Ho, W.K., "A category theoretical argument against the possibility of artificial life: Robert Rosen's central proof revisited", Artificial Life, 12, 1, 117-134, 2006.

[4] Jacob, F., "La logique du vivant », Gallimard, Paris, 1970.

[5] Letelier, J., Marín, G., Mpodozis, J., "Autopoietic and (M,R) systems", Journal of Theoretical Biology, 222, 261–72, 2003.

[6] Letelier, J.C., Soto-Andrade, J., Abarzua, F.G., Cornish-Bowden, A. & Cardenas, M.L., "Organizational invariance and metabolic closure: Analysis in terms of (M,R) systems", Journal of Theoretical Biology, 238(4), 949-961, 2006.

[7] Longo, G., "Critique of Computational Reason in the Natural Sciences", in E. Gelenbe and J.-P. Kahane eds., "Fundamental Concepts in Computer Science", Imperial College Press/World Scientific, 2007.

[8] Louie, A., H., "Any material realization of the (M,R)-systems must have non-computable models", Journal of Integrative Neuroscience, 4,4, 423-436, 2005.

[9] Louie, A., H., "A Living System Must Have Noncomputable Models", Artificial Life, 13, 3, 293-297, 2007.

[10] Maturana, H., & Varela, F.J., "Autopoiesis and cognition: the realization of the living", Reidel, Boston, 1980.

[11] McMullin, B., “Thirty years of computational autopoiesis”, Artificial Life, 10, 277-295, 2004.

[12] Nomura, T., “Category Theoretical Distinction between Autopoiesis and (M,R) Systems”, Proc. 9th European Conference on Artificial Life (ECAL 2007), Lisbona, 2007. (to appear)

[13] Oyama, S., “The ontogeny of information: developmental systems and evolution”, Cambridge University Press, Cambridge, 1985.

[14] Rosen, R., “Life itself: a comprehensive enquiry into the nature, origin and fabrication of life”, Columbia University Press, New York, 1991.

[15] Stewart, J., “La vie existe-t-elle ? Réconcilier la génétique et la biologie”, Vuibert, Paris, 2004.

[16] Wolkenhauer, O., “Interpreting Rosen”, Artificial Life, 13,3, 291-292, 2007.

[17] Zaretzky, A. & Letelier, J., “Metabolic Networks from (M,R) Systems and Autopoiesis Perspective”, J. Bio. Sys., 10 (3), 265-280, 2002.

Motion Sickness in Enactive Human-Computer Systems

T. A. Stoffregen* B. G. Bardy† S. Villard* K. Yoshida* L. Scibora*
(*) Affordance Perception-Action Laboratory, University of Minnesota, USA
(†) Motor Efficiency and Deficiency Laboratory, University of Montpellier-1, France
email: tas@umn.edu, benoit.bardy@univ-montp1.fr, svillard@umn.edu, yoshi028@umn.edu, scibo002@umn.edu

Abstract

Motion sickness (aka cybersickness) is increasingly common in interactive human-computer systems. We review studies motivated by the view that motion sickness may be enactive in nature, that is, that motion sickness may arise from problems of perceptual-motor interactions with the environment and, specifically, may involve subtle, intuitive movements that serve to enhance perception and performance. A central prediction of our approach is that motion sickness will be preceded by unstable control of the body, during exposure to potentially nauseogenic motion. We have confirmed this prediction in laboratory devices, flight simulators, virtual environments, and commercially available video games. We discuss some implications for the design of enactive display systems.

1. Introduction

Motion sickness is an ancient malady. Classically, it occurs with physical displacement (e.g., ships, trains), but recently motion sickness has become a problem with simulated displacement, as in virtual environments. There are widespread reports of motion sickness among users of simulation and virtual environment systems, including systems that depict motion of the user, such as flight and driving simulators, and many virtual environments [17]. The utility of simulation and virtual environment systems, and their acceptance by users, can be limited if they produce motion sickness [1]. This problem provides a practical motivation for understanding the malady. Many virtual environments are visual simulations, and so are considered to create *visually induced motion sickness* (e.g., [8]). Visually induced motion sickness occurs in laboratory devices, such as moving rooms [2; 11; 16; 19], in fixed base flight simulators [17], in projected virtual environments [6], and among players of console video games [12]. There is a qualitative correlation between the technical sophistication of visual simulation and the incidence of motion sickness among users [3, 9]: Generally, the better a system looks, the more likely it is to induce motion sickness. This correlation highlights the importance of behavioral research on the motion sickness in simulator systems: By itself, technological development is making the problem of motion sickness worse, not better. Technological development, *per se*, will not solve this problem. What is needed is behavioural science research relating perception-action to the technological environment.

1.1 Disembodied motion sickness

Classical theories of motion sickness etiology typically have been based on the concept of sensory conflict (e.g., [13]), the idea that motion sickness situations are characterized by patterns of perceptual stimulation that differ from patterns expected on the basis of past experience. Differences between current and expected patterns of perceptual stimulation are interpreted as sensory conflict, which is alleged to produce motion sickness. Despite the intuitive appeal of the sensory conflict concept, theories based on sensory conflict have low predictive validity [4], and may not be scientifically falsifiable (e.g., [5]).

1.2 Embodied motion sickness

The sensory conflict theory of motion sickness is disembodied in the sense that it relies on the comparison of current sensory inputs with inputs expected from a pre-existing, internal model. In embodied theories, causal relations are sought in relations between perception and action, and do not rely on the use (or even the existence) of internal models.

In our research, we investigate a new understanding of motion sickness etiology, which is based on enactive relations between perception and movement, movement and perception. People who are motion sick exhibit unstable control of the body: This is an uncontroversial observation. Our contribution is to suggest that unstable control of the body precedes motion sickness, and that motion sickness occurs only in persons whose control has already become unstable. This is the postural instability theory of motion sickness etiology [15]. We have evaluated the hypothesis that motion sickness will be preceded by unstable control of the body in the context of visually induced motion sickness, and we have confirmed it in a wide variety of contexts, including laboratory devices, operational flight simulators, and console video games. In this presentation, we will review some of the existing research relating motion sickness to perceptually guided movement, and we will discuss implications of this work for the design of enactive technologies.

2. General methodology

Our central hypothesis is that motion sickness will be preceded by unstable control of the body. To test this hypothesis, it is necessary to collect data about motion sickness incidence, and about body movement during exposure to potentially nauseogenic motion. A critical requirement is that researchers know when motion sickness begins, that is, the time at which participants begin to experience the subjective symptoms of motion sickness (e.g., nausea). We examine movement data that were collected before the onset of subjective symptoms of motion sickness, to determine whether there were differences in movement between persons who later became motion sick (the Sick group) and those who did not (the Well group).

One well-established fact about motion sickness is that it is mainly associated with imposed motions in the frequency range 0.1 – 0.4 Hz [7]. In most of our studies, we have used as a motion stimulus optical motion within this frequency range, presented as fore-aft oscillations along the participant's line of sight, or as angular oscillations around the line of sight (Figure 1). We have done this in a moving room (e.g., [2]; Figure 2), in a virtual environment presented via a video projector [6], and in a flight simulator [17]. By contrast, in our research using console video games, we have exercised no control over the motion stimulus: Participants have played "off the shelf" games in the recommended manner. Games have been played either standing, or sitting on a stool.

In all experiments, we have begun by taking baseline measures of motion sickness symptomatology, using the Simulator Sickness Questionnaire (SSQ) [10]. This has been followed by baseline measures of postural activity, with participants standing for 60 s with eyes open, and then with eyes closed, in the absence of any imposed motion stimulus. The main experiment consists of exposure to potentially nauseogenic visual motion stimuli. In studies of console video games, participants have played a given game in a single continuous session for up to 50 minutes. In our other studies, participants have been exposed to visual oscillation stimuli in a series of up to four trials, each 600 s in duration. In all studies, participants are instructed to discontinue participation immediately if they experience any symptoms of motion sickness. At discontinuation or the end of the protocol (whichever comes first), participants fill out the SSQ a second time. We compare pre-exposure and post-exposure SSQ scores for the Sick and Well groups. Movement data are collected using a force platform (yielding kinematics of the Center of Pressure, *COP*), or a magnetic tracking system (yielding kinematics of the head and torso).

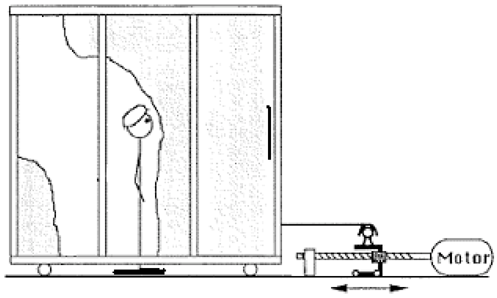


Figure 1. The moving room used in many of our studies.

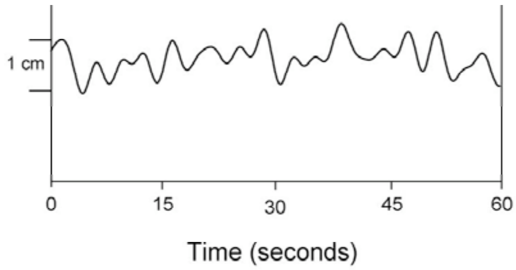


Figure 2. A portion of the sum-of-sines stimulus motion used in many of our studies.

3. Summary of results

In each of our experiments, some (or all) participants have stated that they became motion sick. Across experiments, motion sickness incidence has ranged from 35% (standing in the moving room) to 100% (playing console video games while standing).

In each experiment, post-exposure SSQ scores for the Sick group were significantly higher than pre-exposure scores.

In each experiment, measures of postural activity taken prior to the onset of motion sickness symptoms have exhibited significant differences between Well and Sick groups. Differences have been observed in the COP, the head, and the torso, in such variables as positional variability, velocity, and path length. Each of these parameters is based on the magnitude of motion, and we have found that the Sick group exhibited greater motion magnitude. However, we have also identified Sick/Well differences in parameters of movement that are not based on movement magnitude, such as Detrended Fluctuation Analysis (DFA). In several experiments, scaling exponents for DFA (known as α) have differed between the Sick and Well groups.

We have examined the evolution of postural activity over time during exposure to stimulus motion. To do this, we select three "windows" from the movement data. Each window is two minutes in duration, and the three windows are selected from the beginning, middle, and end of exposure. To ensure equality of exposure, the selection of windows for the Well group has been based on the mean duration of exposure for participants in the Sick group (for details, see [2]). In several experiments, differences between Sick and Well groups have been found to develop over time. Differential changes in movement over time have been observed in classical variables (Figure 3), but also in DFA (Figure 4).

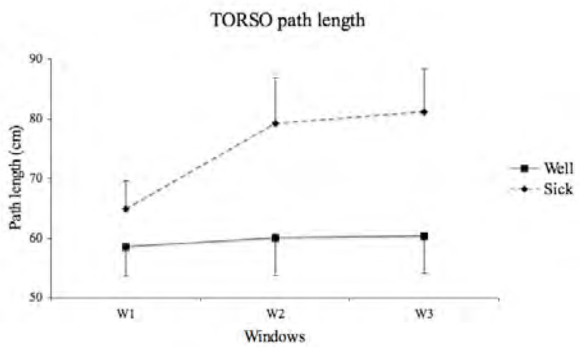


Figure 3. The linear path length of torso movement changes differentially over time (during exposure) for Well and Sick groups.

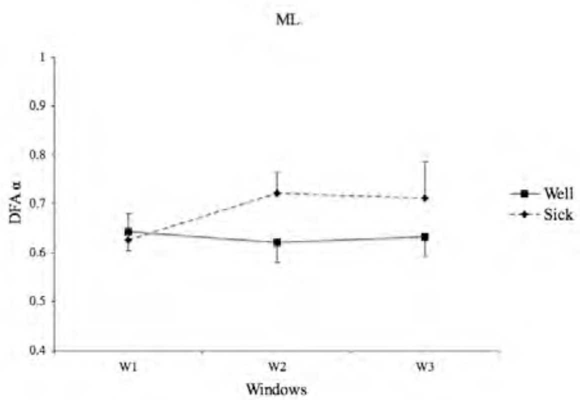


Figure 4. The scaling exponent, α, obtained from DFA, changes differentially over time (during exposure) for Well and Sick groups.

We have used data on body movement to develop mathematical models that can be used to predict motion sickness [16]. Our models are unique in that they include only movement terms (i.e., no terms relating to perception or expectation), and because they account for up to 60% of the variance in actual motion sickness incidence; models that do not include movement typically account for no more than 35% of the variance.

4. Discussion

In a variety of studies conducted across a range of venues, we have shown that the subjective phenomena of visually induced motion sickness occur subsequent to changes in body movement. Our predictions, and many of our manipulations are not motivated by the sensory conflict theory of motion sickness [13], but are directly motivated by the postural instability theory of motion sickness [15], and the results of our studies consistently support the postural instability theory.

Postural activity serves to maintain the body's orientation relative to ambient forces (i.e., the gravito-inertial force environment). In addition, postural activity can simultaneously support the achievement of perception-action goals that are "supra-postural", such as reading [14]. Recent research has confirmed that postural activity during unperturbed stance is modulated by variations in the perceptual demand of supra-postural tasks [e.g., 18]. In our view, postural control is routinely integrated with perceptual activity (e.g., eye movements) to optimize performance. It may be that complex imposed motion (e.g., ship motion) not only destabilizes overall body orientation, but also destabilizes the functional linkage between body sway and perceptual performance. Such an effect would strengthen our claim that motion sickness is a disruption of enaction.

Our focus on visually induced motion sickness means that our research may have direct relevance to motion sickness or cybersickness that is associated with visual displays and other types of visual virtual environments.

A challenge for our approach will be to extend our empirical assessments of relations between intuitive movement and motion sickness to include multi-sensory display systems, as well as motion sickness that follows inertial displacement, such as sea sickness.

An additional challenge will be to increase the relevance of our work to technologies that are specifically enactive in intent. One way to do this could be to vary participants' tasks in interaction with our displays, so as to vary the relevance of intuitive postural activity to task performance. We would expect that postural instability and subsequent motion sickness would be more likely to occur when imposed motion stimuli are relevant to tasks performed in the context of a virtual environment, for example, when participants attempted to move deliberately relative to a motion stimulus, rather than merely observing it. Such a prediction, if confirmed, could imply that enactive interface technologies may be especially sensitive to issues of perceptual-motor stability of users.

References

[1] Biocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environment technology? *Presence*, *1*, 334-343.

[2] Bonnet, C. T., Faugloire, E. M., Riley, M. A., Bardy, B. G., & Stoffregen, T. A. (2006). Motion sickness preceded by unstable displacements of the center of pressure. *Human Movement Science*, *25*, 800-820, Champaign IL.

[3] Crowley, J. S. (1987). Simulator sickness: A problem for Army aviation. *Aviation, Space, and Environmental. Medicine, 58*, 355-357.

[4] Draper, M. H., Viirre, E. S., Gawron, V. J., & Furness, T. A. (2001). The effects of virutal image scale and system delay on simulator sickness within head-coupled virtual environments. *Human Factors*, *43*, San Diego CA.

[5] Ebenholtz, S. M., Cohen, M. M., & Linder, B. J. The possible role of nystagmus in motion sickness: A hypothesis. (1994). *Aviation, Space, & Environmental Medicine*, *65*, 1032-1035.

[6] Flanagan, M. B., Albanese, G., Villard, S., & Stoffregen, T. A. (2007). Postural instability and motion sickness in a virtual moving room. Manuscript submitted for publication.

[7] Guignard, J. C., & McCauley, M. E. (1990). The accelerative stimulus for motion sickness. In G. H. Crampton (Ed.), *Motion and space sickness* (pp. 123-152). Boca Raton, Fl: CRC Press.

[8] Hettinger, L. J.; Riccio, G. E. (1992). Visually-induced motion sickness in virtual environments. *Presence*, *1*, 306-310.

[9] Kennedy, R. S., & Fowlkes, J. E. (1992). Simulator sickness is polygenic and polysymptomatic: Implications for research. *International Journal of Aviation Psychology*, *2*, 23-38.

[10] Kennedy, R. S., & Lane, N. E. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology, 3*, 203-220.

[11] Lishman, J. R., & Lee, D. N. (1973). The autonomy of visual kineasethesis. *Perception, 2*, 287-294.

[12] Merhi, O., Faugloire, E., Flanagan, M., & Stoffregen, T. A. (2007). Motion sickness, console video games, and head mounted displays. *Human Factors*, *49*, 920-934, San Diego CA.

[13] Reason, J. T. (1978). Motion sickness adaptation: A neural mismatch model. *Proceedings of the Royal Society of Medicine, 71*, 819-829.

[14] Riccio, G. E., & Stoffregen, T. A. (1988). Affordances as constraints on the control of stance. *Human Movement Science, 7,* 265-300, Champaign IL.

[15] Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. *Ecological Psychology, 3,* 195-240, Mahwah NJ, 1991.

[16] Smart, L. J., Stoffregen, T. A., & Bardy, B. G. (2002). Visually-induced motion sickness predicted by postural instability. *Human Factors, 44,* 451-465, San Diego CA.

[17] Stoffregen, T. A., Hettinger, L. J., Haas, M. W., Roe, M., & Smart, L. J. (2000). Postural instability and motion sickness in a fixed-base flight simulator. *Human Factors*, *42,* 458-469, San Diego CA.

[18] Stoffregen, T. A., Hove, P., Bardy, B. G., Riley, M. A., Bonnet, C. T. (2007). Postural stabilization of perceptual but not cognitive performance. *Journal of Motor Behavior*, *39*, 126-138, Champaign IL.

[19] Stoffregen, T. A., & Smart, L. J. (1998). Postural instability precedes motion sickness. *Brain Research Bulletin*, *47*, 437-448, London.

Tests of Applications Intended to Increase Non-visual Access of Web Content

Chui Chui Tan* Charlotte Magnusson *† Wai Yu*

(*)Queen's University Belfast, UK

(†)Lund University, Lund, Sweden*

E-mail: s.tan@qub.ac.uk, charlotte.magnusson@certec.lth.se, w.yu@qub.ac.uk

Abstract

This paper discusses the user tests carried out for the launching interface of a Web graphic-based system by using either a keyboard or the PHANToM. This interface is designed to provide an accessible end interface for visually impaired people to gain access to different types of graphics on the Web. Two different launching interfaces have been developed with the same objective and they have been tested with the visually impaired users. The results of the tests are discussed and compared in the paper. They revealed that both methods are useful for visually impaired people to gain access to Web graphics, such as a 3D content.

1. Introduction

Research on presenting Web content through non-visual form is very active especially after the GUI interface has become one of the most prevalent interface types which exist today. Haptic, auditory and multimodal modalities have been used to provide novel methods for people with visually impairments to interact with computer systems. For instance, Ramstein et al [1] have proposed the PC-Access system that offers auditory information, reinforced by the sense of touch to enable blind users to be aware of what happens on the screen and where the pointer is. Rosenberg and Scott [2] have also carried out a study to prove that force feedback can enhance a user's ability to perform basic functions within graphical user interfaces.

The work described in this paper is part of a European project, ENABLED. One of the objectives of the *ENABLED* system is to provide an accessible end interface for people with vision difficulties to gain access to various types of graphics on the Web [3]. A unique aspect of this system is that it allows users to select a graphical content for further exploration directly from the Web page. The first version of the system design involves two different ways of allowing users to select and access graphics, for example a 3D content. The first is through a button placed on a web page (called the *ENABLED* button, see figure 1 to the left) and the second is through a MSAA based haptic interface to the web browser [3], [4] (see figure 1 and 2). For the former approach, an *ENABLED* button which associates with each graphical content on the Web page is used as a trigger switch. By selecting the button for the graphics and pressing the *ENTER* key, a sub-application will be launched to present an accessible interface for the user to access to the graphics via haptic and audio feedback.

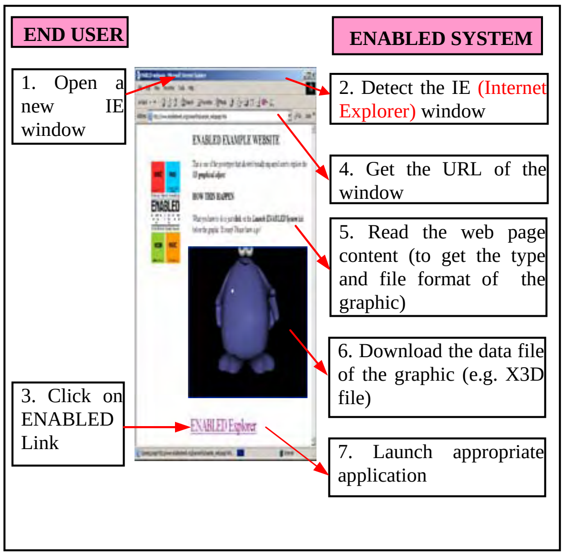


Figure 1. The ENABLED button approach.

In the second approach, the browser is represented as a haptic box on which link positions are indicated by a decrease in stiffness combined with horizontal forces which generate a "fake" depression over the link [5]. Audio feedback is generated both by the use of ZoomText and by a weak white noise in stereo (the stereo position reflected the user position on the browser) which grows louder when the user is on a link. To follow a link the user clicks on the button on the PHANToM stylus, and in the case the user clicks on a X3D file the application for displaying the content of the file haptically is launched.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

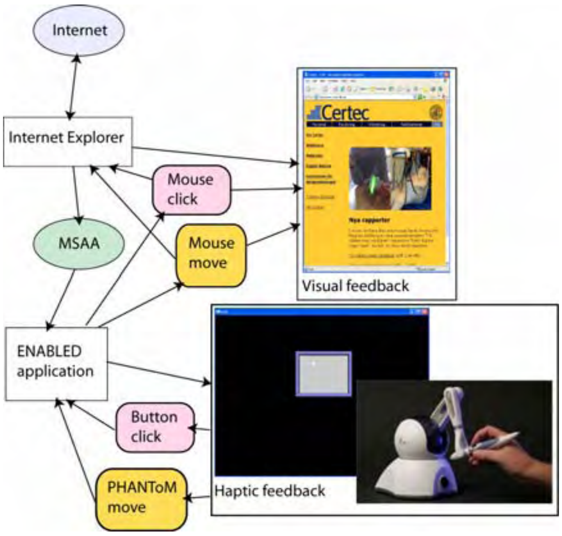


Figure 2. The haptic browser.

The current paper reports on the outcome of two explorative user studies performed to obtain user feedback and to ensure user involvement in the design process.

2. Tests of the ENABLED button (int the UK)

Six blind participants were recruited for the test (four congenitally blind and the others are blind after birth; all are screen reader - JAWS users). The participants were firstly introduced to the interface and the functionality of the ENABLED button. Then they were asked to navigate a Web page which contained four different graphics (including, two graphs, a map and a 3D content) in accordance with JAWS. The task required the participants to search for the ENABLED button for a specific graphical content, which is the 3D Content. They were requested to launch the sub-application by using the ENABLED button as described in the previous section. After carrying out the tests, the participants were asked to indicate their thoughts about the interface and the problems they encounter during the test.

Overall participants' views on the launching interface through the ENABLED button were positive, and all considered that the concept was sound. The interface was intuitive and they were satisfied with the features of the button. However one of the participants commented that the location of the ENABLED buttons should be standardized and positioned under the graphics they are associated to. In general, participants were pleased that the selection and launching of the sub-application could be performed solely with a keyboard, which is their most commonly used input device. Additionally, they were gratified that this system could be used directly from the Web page.

3. Test of the haptic browser (in Sweden)

The test in Lund was performed as an informal exploratory test, where four participants navigated within a web browser and then explored a 3D object (a X3D file) placed on a page. For the ENABLED button approach, the participants were explained about the functionality of the button and were instructed to put the focus on the appropriate button and activate it using JAWS. They were then asked to explore the resulting 3D object. In the test of the haptic browser, the functionality of the browser was first demonstrated, and afterwards the participant was asked to browse around for a while freely. After some time the test leader navigated to the page with a 3D object link and asked the participant to click this link. Finally the participant was asked to explore the 3D object that was displayed as a result of this action by using the PHANToM Omni.

The results of this test revealed that it is possible for blind and visually impaired users to use a haptic browser as described in [4]. The haptic feedback on the links was appreciated, though it became obvious that the type of feedback given (which reflected the visual shape of the link) made it easy to find links organized in a column, but quite hard to find links in a row. The audio feedback was commented as "annoying" and therefore re-design is needed. When asked about preferences, most participants prefer the ENABLED button version (while wanting the haptic alternative as a complement).

For the 3D information it was seen as useful, and the type of information wanted is exemplified as excel charts, maps, diagrams, 3D models of famous buildings, strange shapes, to get an overview, medical teaching material, understanding a layout and knowing where buttons and form fields are. For the fairly simple 3D model used in the test it was possible for the users to zoom in and out despite the crude zooming function implemented. It is expected that a more elaborate zoom function which preserves the relative position of the user with respect to the model will be needed for more complex 3D models. The test further highlighted the need for labels on the parts of a figure (like "arm", "leg" etc). A somewhat surprising result was the fact that one user preferred not to have a limiting box (in contrast to what is usually recommended [1], [6]). This indicates that the design of such a box needs to be further investigated.

4. Conclusion

These initial tests show that the design of the *ENABLED* system makes it possible to access graphical and 3D content on the Web via the use of a keyboard and the PHANToM. In general, users prefer to use an existing screen reader accompanied by a specially designed button ("the *ENABLED* button") to access graphics on the Web, though a haptic browser was seen to provide a useful complement. However, the first approach is of advantage for those who do not have experience on haptic and own only the keyboard, and wish to directly approach to graphics. Whereas the second approach is useful when the user has a PHANToM and would like to explore more about the Web page, for example the spatial layout and presentation of the Web page. The overall system gives a great opportunity for visually impaired people to gain access to graphical Web content via haptic and audio modalities.

5. Acknowledgement

This work is supported by ENABLED - an Integrated Project funded by the EU Framework 6 Programme, IST (FP6-2003-IST-2-004778).

References

[1] C. Ramstein, O. Martial, A. Dufresne, M. Carignan, P. Chasse, and P. Mabilleau, "Touching and Hearing GUI's: Design Issues for the PC-Access System", ACM ASSETS 1996, ACM Press, pp. 2-9.

[2] L. Rosenberg, and S. Scott, "Using Force Feedback to Enhance Human Performance in Graphical User Interfaces", Proc. CHI 1996, ACM Press, pp. 291-292.

[3] C.C. Tan, W. Yu, and G. McAllister, "Developing an ENABLED Adaptive Architecture to Enhance Internet Accessibility for Visually Impaired People", Proc. ICDVRAT 2006, pp. 231-238.

[4] C. Magnusson, C.C. Tan, W. Yu, "Haptic Access to 3D Objects on the Web", Proc. EuroHaptics 2006.

[5] G. Robles-De-La-Torre, V. Hayward, 2001, "Force Can Overcome Object Geometry, In The Perception Of Shape Through Active Touch", Nature, Vol. 412, pp. 445-448.

[6] Miller, T. and Zelesnik, R. An Insidious Haptic Invasion. Proc UIST, (1998), 59-64.

Relationship between ego- and exo-centric judgments: influence of somaesthetic cues

Trousselard Marion* Ceyte Hadrien* Cian Corinne*
(*) Centre de Recherches du Service de Santé des Armées, La Tronche cedex,
E-mail: mtrousselard@crssa.net, hceyte@crssa.net, corinnecian@crssa.net

Abstract

Interactions between organisms and surfaces of support constitute a major factor contributing to the control of orientation. In particular, the tactile system which conveys relevant gravitational information increases feedback information systems. A major issue in orientation control studies is to determine whether and how we use spatial frames of reference to organize our spatially oriented behaviours. The aim of the present study was to evaluate the role of tactile cues in the ego- and exo-centric frames of reference perception as in their relationship.

Four experiments assessed the importance of the tactile information on exo-centric (Subjective Visual Vertical (SVV), body orientation (BO)) and egocentric (Z-axis orientation) static perception. Whether this somaesthetic information contributed to all perceptual settings, it appeared that egocentric judgment could constitute one basis of the oriented behavioral organization towards the extra-corporal.

1. Contribution of somaesthetic cues to the perception of body orientation and subjective visual vertical

Without relevant visual cues, the visual perception of the true vertical (SVV) is biased in roll tilted subjects toward the body axis. Though it is often argued that this so-called Aubert effect originates in an underestimation of one's own body position, there is no convincing experimental evidence that visual and postural orientation are directly related [9].

The determination of body orientation and the SVV are considered to be a consequence of changes in the vestibular and somatosensory inputs secondary to a body tilt [2]. The otolithic system is a source of information for the orientation of the head with regard to the direction of gravity. The somaesthetic system is assumed to provide information about body orientation, notably in response to the anti-gravitational forces with an increasing role with gradually body tilts.

Experiment 1 investigated the somaesthetic system by means of a body-cast. This technology induces smaller asymmetries and a more diffuse/homogenous repartition of body pressure when tilting whereas gravity-based otolithic information is still present.

To refer to the explanation of the Aubert effect as a consequence of an underestimated body orientation, we expected an increased Aubert effect when the available somatosensory inputs became diffuse by the body-cast because of a greater underestimation of the body position in the body-cast.

1.1. Method and results

In otherwise dark room, 16 naive subjects, seated in a tilt-chair capable of rotating to any angle in all planes of space, were rolled sideways from 0 deg to 105 deg ($v=3$ deg. s^{-1}, $a=.01$ deg.s^{-2}). At the pre-set tilt, subjects were asked (1) to verbally indicate, eyes closed, the subjective body orientation using a clock scale and (2) to adjust a luminous line to the vertical under two conditions. In the strapped condition, subjects were restrained by means of straps; In the body cast condition subjects were completely immobilized by means of the body-cast (Figure 1).

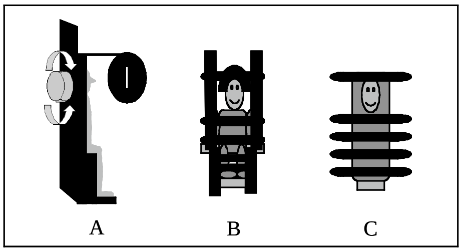


Figure 1. Tilt chair (A), Strapped (B) and body-cast (C) conditions

Data analyses were carried out on the angular difference (in degrees) between (1) the estimated body orientation and chair tilt (positive for the overestimation, negative for the underestimation), (2) the rod adjustment and the upright (positive for Aubert effect).

Results showed a greater A-effect (Figure 2; * denotes a significant difference between conditions $p < .05$) but an overestimation of the body orientation (Figure 3; $p < .05$) in the body cast condition for the higher tilt values (beyond 60º).

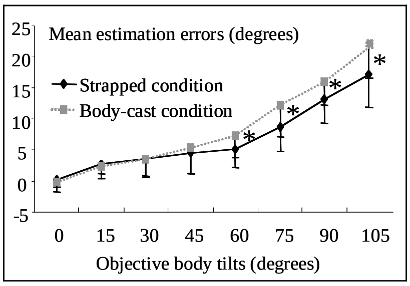


Figure 2. Subjective Visual Vertical under strapped and body-cast conditions.

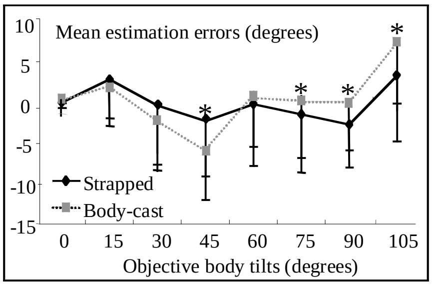


Figure 3. Perceived body orientation under strapped and body-cast conditions.

Visual and postural errors were not directly related (no correlation). However, above 45 deg, the body-cast condition produced the same response shift in the same direction for both SVV and body orientation estimations (roughly 5 deg).

1.2. Discussion

Since the otolith organs respond to the gravity in the same way in both conditions, errors in body orientation estimation and visual vertical are due to the change in somaesthetic cues. However, conversely to the hypothesis, (1) the body orientation estimation did not account for the visual vertical errors and (2) the absence of relationship between them suggested separate mechanisms for the two perceptual tasks under body tilt.

One explanation of the body-cast effect focuses on the process of integration of the available informations. The patterns of pressure distribution and muscle tone are unusual whereas gravity-based otolithic cues do not change. It has been suggested that the missing congruence between sensory cues raises the question that «spatial disorientation » per se may interfere with thinking processes involved in perceived body orientation and SVV [4].

Nevertheless, in the body-cast condition, the same response shift occurred in the same direction for both the perceived visual vertical and body orientation. Thus, in the absence of relevant visual information, subjects use the available somaesthetic cues to provide an estimate of orientation of body-in-space, that serve both as a basis for the self estimates and is used in the estimates of subjective visual orientation with respect to earth co-ordinates.

2. Contribution of somaesthetic cues to the perception of Z-axis

Body midline constitutes one basis of the oriented behavioral organization towards the extra-corporal space [7]. It can be defined as the plane composed by the orthogonal projection of the Z-axis (virtual head-to-foot axis) on the frontal plane. Subject-relative information holds its relation to the body regardless of its orientation in earth co-ordinates. However, it has been shown that this egocentric judgment is disturbed when the orientation of the Z-axis is no more parallel to the direction of gravity [1, 8].

Few researches have directly investigated the effect of body tilt on the perception of Z-axis, and it remains unclear whether this egocentric reference needs to be encoded in a gravity-based frame of reference or whether it is the action of the gravitational forces on the different sensory systems which induces a displacement of the equilibrium position between information arising from both sides of the body space [5]. Thus, Experiments 2 and 3 aimed at reassessing the influence of gravity direction on the perceived Z-axis.

2.1. General method

For the two following experiments, 7 naive's participants sat in the tilt-chair (Figure 1 A). The orientation of the body midline (perceived Z-axis) was investigated by the rolling adjustment of a rod on the subjects' Z-axis.

Data analyses were carried out on the angular difference (in degrees) between the true orientation of the Z-axis and participants' response (positive if the rod deviated beyond the objective body orientation (errors in the direction of body tilt) and negative if the rod deviated below the objective body orientation (errors in the opposite direction)).

2.2. Role of the gravitational cues when the Z-axis and gravity direction are no more aligned

Experiment 2 investigated the effect of a dissociation between the Z-axis and the direction of gravity by placing subjects in roll and pitch tilt postures. Lying horizontally on the side (roll tilt) or in a supine posture (pitch tilt) induced a 90 deg modification of the direction of the Z-axis relative to gravity. However, the two body tilts did not allow processing the potential sensory information about the direction of gravity in the same way. Unlike roll tilt posture, supine body orientation renders the otholitic and somaesthetic systems irrelevant to the direction of gravity [8]. It was hypothesized that the perception of the Z-axis is only influenced by the presence of direct gravitational information (roll tilt condition).

Results showed that the perceived Z-axis deviated in the direction of body tilt for the roll tilt posture (p<.05) whereas the pitch body orientation did not alter the Z-axis judgment (Figure 4).

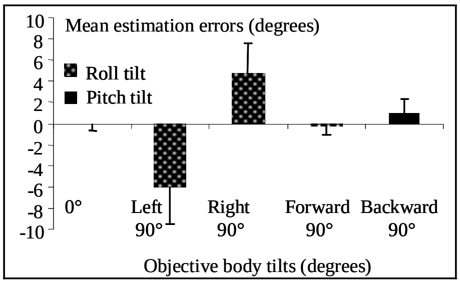


Figure 4. Perceived Z-axis orientation under 90 degrees of roll and pitch tilt postures

These findings suggested that the egocentric perception of the body is encoded in a gravity-based system of reference only when direct information about the direction of gravity is present.

2.3. Role of the somaesthetic cues in the perceived Z-axis

Experiment 3 further determined the relative importance of the otholitic and somaesthetic systems in the perceived Z-axis by reducing gravity-based somaesthetic cues in roll tilt postures (body-cast).

The changes in the available somaesthetic information induced by the body-cast could induce less Z-axis errors than for the strapped condition.

In the body-cast, results showed errors deviating more in the direction of body tilt (Figure 5; p<.05). Moreover, an absence of correlation was observed between errors in the two restriction conditions.

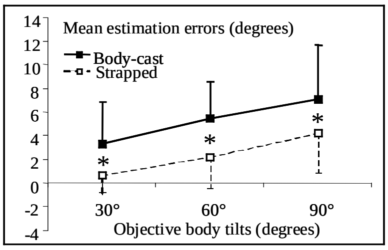


Figure 5. Perceived Z-axis orientation under strapped and body-cast conditions.

Contrary to our hypothesis, adjusting a rod on the orientation of Z-axis produced errors in the direction of body tilt greater in the body cast than in the strapped condition. Since the otoliths produced the same gravity response in the two conditions, somaesthetic inputs appeared necessary for the estimation of object orientation with respect to the body.

However, the increasing errors in the body-cast did not allow rejecting the hypothetical displacement of the equilibrium position between information arising from both sides of the body in the strapped condition. A sensory discordance relative to the representation of the body in space could be created by the body-cast between tactile cues (information about body orientation) and otolithic inputs (information about head orientation). As gravity-based somaesthetic cues were greatly reduced in the body-cast, the head and the body may perceive differently a same tilt with respect to the direction of gravity. The body-cast could imply a dissociation of the perceived body and head orientation inducing a greater misperception of one's own body.

2.4. Discussion

Whether the present research showed that egocentric judgments depend on gravitational information, the processes which give rise to ego-centric body co-ordinates kept unknown [3, 5]. Overall, the present experiments suggested that there is no global or absolute representation of the ego-centric space which seems continuously updated, depending on the characteristics of the situation

3. Influence of the perceived ego-centric co-ordinates on the SVV

Faced to the classical assumption explaining the Aubert effect by an underestimation of one’s own body orientation, it can be suggested that SVV errors could also be derived from the ego-centric processing. It is reasonable to think that once subjects perceived their body orientation relative to the gravity, they have to realize an angle between their Z-axis and the rod according to the perceived body orientation. Thus the misperception of Z-axis as the origin of the produced angle can produce misjudgments of the visual vertical.

The goal of Experiment 4 was to investigate whether the misperception of the Z-axis [3] could account for errors in the visual vertical.

3.1. Method and results

11 naïve subjects were restrained in the tilting chair to perform SVV and Z-axis tasks in two separate sessions. Protocols, data analyses and conventions were described in Experiments 1, 2 and 3.

Results showed that roll tilts induced increasing errors with body tilt in the direction of body tilt for both the perceived Z-axis and the SVV (Figure 6; $p<.05$). Moreover, subjects with the higher Aubert effect exhibited the higher shift of the perceived Z-axis (positive correlation; $p<.05$).

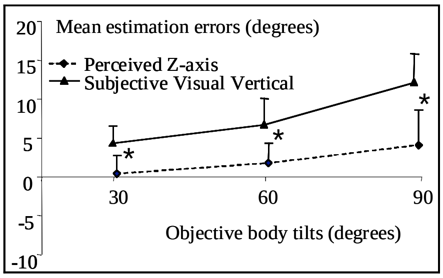


Figure 6. Perceived Z-axis and Subjective Visual Vertical.

A linear regression analysis applied to the mean individual errors in the three body tilts for each task showed that (1) all estimates seemed to be shifted in the direction of body tilt (positive slope) and (2) the two straight lines were parallel [deviation from parallelism $p>.05$], but separated by an offset.

3.2. Discussion

When Z-axis and gravity directions were not aligned, a roll body tilt induced correlated errors in the exo-centric and ego-centric orientation assessments. The more the perceived Z-axis is shifted in the direction of body tilt, the higher the Aubert effect, suggesting that SVV adjustments will not be realized from the objective Z-axis but from an orientation where the rod seems parallel to it (perceived Z-axis). However, the ego-centric estimates cannot fully account for the SVV errors.

4. Conclusions

This present research explored the role of tactile cues on the perception of spatial frames of reference. It was observed that (1) a misperception of the Z-axis could partly account for the misperception of the direction of gravity and (2) changes in the tactile cues following a body tilt are implied in both ego and exo-centric static perception.

Thus, the ego-centric Z-axis reference seems to constitute one basis of the oriented behavioral organization towards the extra-corporal space. Further research on the construction of subjects' internal body representation, and on the underlying mechanisms involved in its processing, are valuable to understand the role of the ego-centric reference in the spatially oriented behaviors.

References

[1] Bauermeister, M., "Effect of body tilt on apparent verticality, apparent body position, and their relation", Journal of Experimental Psychology, 67(2), 142-147. (1964).

[2] Bronstein, A.M., "The interaction of otolith and proprioceptive information in the perception of verticality. The effects of labyrinthine and CNS diseases", Annals of the New York Academy of Sciences, 871, 324-333, 1999.

[3] Ceyte, H., Cian, C., Nougier, V., Olivier, I., and Trousselard, M. "Role of gravity-based information on the orientation and localization of the perceived body midline", Experimental Brain research, 176, 504-509, 2007.

[4] Ito, Y, and Gresty, M.A., "Subjective postural orientation and visual vertical during slow pitch tilt for the seated human subject", Aviation, Space, and Environmental Medicine, 68, 3-12, 1997.

[5] Jeannerod, M., and Biguer, B., "The directional coding of reaching movements. A visuomotor conception of spatial neglect", Jeannerod Ed, Elsevier Science North-Holland, Amsterdam, 1987.

[6] Mast, F.W., and Jarchow, T., "Perceived body position and the visual horizontal", Brain Research Bulletin, 40(5-6), 393-397, 1996.

[7] Prieur J.M., Bourdin, C., Sarès, F. and Vercher, J.L., "Effects of head orientation and lateral body tilt on egocentric coding: cognitive and sensory-motor accuracy", Journal of vestibular Research, 16, 93-103, 2006.

[8] Templeton, W.B., "The role of gravitational cues in the judgment of visual orientation", Perception and Psychophysics, 14(3), 451-457, 1973.

[9] Van Beuzekom, A.D., and Van Gisbergen, J.A.M., "Properties of the internal representation of gravity inferred from spatial-direction and body-tilt estimates", Journal of Physiology, 84, 11-27, 2000.

Haptic Feedback in Large-Scale VEs: Evaluation of SPIDAR-G

Sehat Ullah[*] Samir Otmane [†] Paul Richard [‡]*
([*]) Laboratoire IBISC, 40 rue du Pelvoux, 91000 Evry, France
([†]) Laboratoire IBISC, 40 rue du Pelvoux, 91000 Evry, France*
([‡]) Laboratoire LISA, 62 avenue Notre-dame du Lac, Angers, France
Email: Sehat. ullah@ibisc.univ-evry.fr, Samir.Otmane@ibisc.univ-evry.fr, richard@istia.univ-angers.fr

Abstract

This paper deals with haptic feedback in large-scale Virtual Environments (VEs). After a short survey and comparison of haptic devices, we present an original integration of SPIDAR-G. The device position measuring capacity is given. A VE where the user can select and manipulate virtual objects while feeling their weight and resistance offered by a static obstacle was developed. A preliminary evaluation of the system is presented. Results from questionnaire revealed a very good acceptance and feeling from the users.

1. Introduction

To increase the user's immersion in Virtual Environments (VEs) and to impart them with realistic experience, several sensory modalities like 3D vision and sound have been added with success.

The use of these modalities along with high quality interactive scenes resulted in VEs that allow more comprehensive understanding of problems.

However to fully understand VEs, to make them more realistic and to increase human performance, the haptic feeling of their physical characteristics (inertia, weight and hardness) is required. Thus the inclusion of haptic modalities becomes more compulsory.

In this paper, we present an original integration and a human performance evaluation of SPIDAR-G in a large-scale VEs. The user can select and manipulate virtual objects while feeling their weight and resistance offered by a static obstacle.

The paper is organized in the following manner. In section 2 we present a short survey of existing haptic interfaces and we make a comparison between SPIDAR [1, 2] and other force feedback devices. Section 3 describes the SPIDAR-GL (Large-scale SPIDAR G). Section 4 describes the VE developed for experimentation with the SPIDAR-GL. Section 5 presents a preliminary user study.

2. Related work

In order to provide force feedback in VEs, a number of technological solutions have been proposed and utilized. For example the PHANTOM™ desktop has a (160 x 120 x 120) mm^3 workspace and returns force on a single finger [3]. Some other models of PHANTOM™ and their characteristics are present in table1 [4].

Table 1. Characteristics of PHANTOMTM devices.

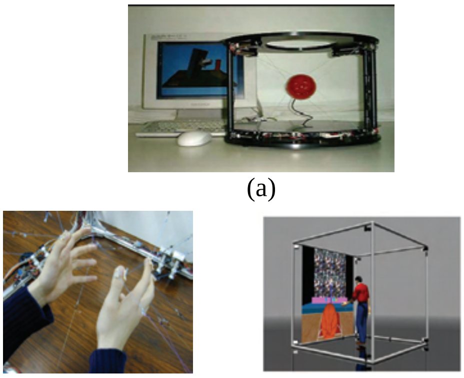
Model	DOF (output)	Workspace (in)	Maximum Force (N)
PHANTOM 1.5	3/6	381x267x191	8.5
PHANTOM 3.0	3/6	838x584x406	22
PHANTOM Premium	3	381x267x191	8.5

Such devices normally suffer from small workspace and integration problems with immersive displays.

Alternative solutions provide finger force feedback to the user. For example, the CyberGrasp™ can exert a force of 12N on each finger of the user's hand. It has a workspace of 1m-radius hemisphere and the user has to bear its weight of 450g.

The CyberForce™ that provides external force feedback [5] is very precise in measuring position and orientation of the user's hand and can prevent penetration into a simulated wall. However, it offers a small workspace of 12"x 12".

Sato proposed a more interesting and less intrusive approach, based on the use of strings. The resulting devices are called SPIDAR (Space Interface Devices for Artificial Reality). In [6] two SPIDAR-G systems were used for object manipulation in a small-size desktop VE. Task completion time was analyzed under different conditions including haptic feedback. In [7] a SPIDAR-H was used in a human scale VE for product design and maintenance tasks. In this study, the SPIDAR was only used to provide position tracking and the feeling of objects weight.



(a)

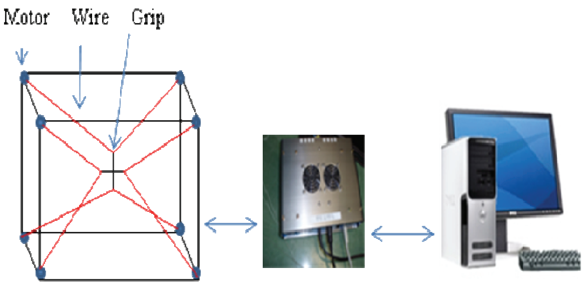
(b) (c)

Figure 1. Illustration of different SPIDAR:
(a) SPIDAR-G, (b) SPIDAR-8, (c) SPIDAR-H

3. System description

3.1 SPIDAR-GL

We have mounted the 8 motors of SPIDAR-GL on the corners of a large size (1.8 x 1.2 x 1.2) m^3 iron frame. All these motors are connected to the HDHC (High Definition Haptic Controller) that communicate with PC via USB 2.0. We placed this frame in front of a (1.2m x.9m) screen. The system provides 6DOF, with translation and Orientation force and is capable to measure a distance of .80m, .50m and .50m in X, Y and Z directions respectively from the origin (center of frame). An illustration of the system configuration and a picture of the system is given in the figure 2 and figure 3 respectively.



(a) (b) (c)

Figure 2. System configuration: (a) frame and strings, (b) HDHC controller, (c) Computer.

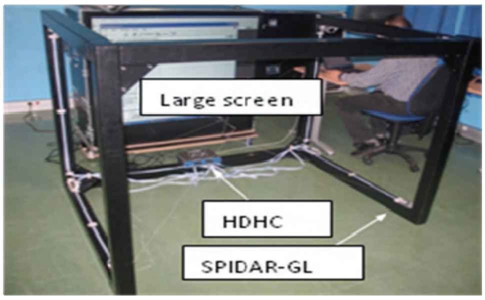


Figure 3. SPIDAR-GL of IBISC lab.

3.2 System calibration

In order to calibrate the SPIDAR we took the centre of the frame as home position (i.e. origin) and calculated the coordinates for both ends (i.e. one end attached to the grip and the other attached to the pulley) of each string.

The coordinates of strings were stored in a text file (called device information file). Similarly the phase sequence of each motor was set on hit and trial bases and was also stored in this file.

Our program developed in C++, which communicates with HDHC controller, makes use of the calibration routine that comes as a part of the SPIDAR API. The calibration routine takes the device information file as input and gives the initial position and orientation of the grip as output. Before the execution of the calibration routine the user is required to hold the grip at the middle of the frame. Figure 4 presents an illustration of the calibration process, where P represents position and O represents orientation of the grip.

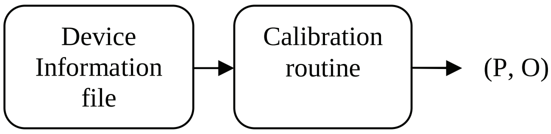


Figure 4. Illustration of the calibration method.

4. Virtual Environment

In order to carried out a usability study with our system we developed the VE that is given in figure 5.

Figure 5. Experimental virtual environment

The VE is a virtual square room that contains a total of eight cubes. Three walls (left, right and back) and a screen act as static obstacles. A wood-textured floor represents the limit of the room in the downward direction. A small black sphere is used as a pointer and follows the movements of the SPIDAR-G grip.

The virtual cubes are made from the same material but have different size. Each cube is associated with a mass ranging from 200g (smaller cube) to 1kg (bigger cube). The user can select any cube using the pointer. For selection, collision detection between a cube and the pointer is used as a metaphor. When selected, the corresponding cube is attached to the pointer and follows its movements. Thus any cube can be moved and placed on any of the red or green zone. As the SPIDAR has no binary input for selection/release confirmation, we have used a metaphor for the placement of the cubes on the two zones. This metaphor is based on the collision detection between a zone and a cube.

As the user selects a cube using the pointer, he feels the its weight. For each cube the weight is calculated using the following the formula:

$$W_t = m_o g + W_c$$

Where W_t represents the total weight of a cube. m_o is the mass associated to a cube and g represents the gravitational force. W_c is the weight of the grip which is added to the weight of each object.

5. Software architecture

Our software that enabled us to use the SPIDAR in the VE has client-server architecture as illustrated in the figure 6.

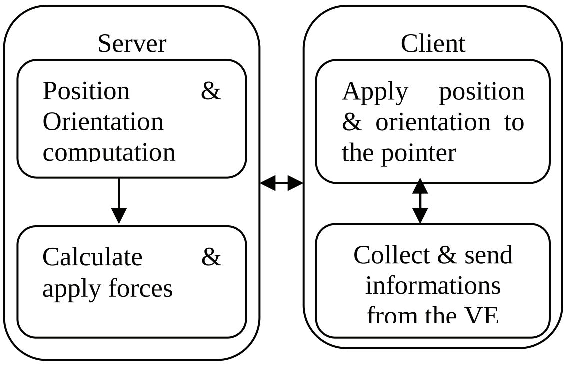


Figure 6. Illustration of software architecture.

We developed the server part of this software using C++ language. This part of the software performs the following tasks:

1. Establish PC and HDHC controller communication
2. Take the calculated position and orientation from the HDHC controller.
3. Establish connection with Virtools client.
4. Calculate and apply the force or weight based on the information received from the client.

The client part of this software was developed using Virtools Dev4.0 environment. This part is responsible for the presentation of virtual environment and supports the interactivity of the virtual objects. The position and orientation send by the SPIDAR server are applied to the pointer. Based on the collision detection mechanism, object selection is checked. The selected object is attached to the pointer and its ID is send to the server for force calculation. Similarly the collision is checked for object placement in a particular zone. In addition the client detects collision between the selected cube and the static obstacles (walls or screen) and the corresponding information is sent to the SPIDAR server for creating a resistive force in a particular direction.

By default the Virtools Dev. software has no support for communication with SPIDAR. In order to establish this communication and to enable the integration of the SPIDAR device in Virtools Dev. environment, we developed three building blocks (BB). These building blocks are presented in figure 7.

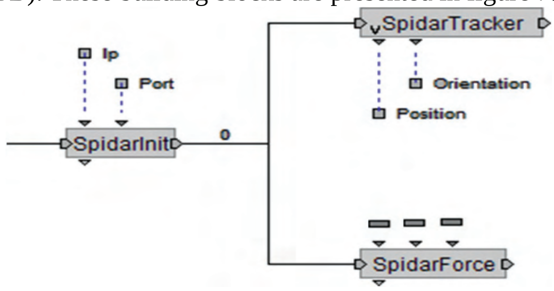


Figure 7. Building Blocks developed for SPIDAR to be used in Virtools.

The BB **SpidarInit** is used for communication and data exchange between client and server. **SpidarTracker** BB takes the position and orientation and forwards it to another BB for application to the pointer. The BB **SpidarForce** takes the data from the rest of the program about object selection, releasing and collision with obstacles.

We tested this software on two different system configurations: (1) firstly, we installed both SpidarServer and Virtools client on the same machine and then each on a different machine.

6. System evaluation

In order to evaluate our system, we carried out a usability study. A total of 13 young volunteer students including 11 males and 2 females participated in the experiment. All participants performed without any knowledge of the system. The objective of this experiment was twofold: evaluate (1) the overall usability of the SPIDAR-G in tasks involving manipulation of virtual objects with force feedback, and (2) the subjective effect of haptic cues on user performance. Therefore, each user performed the experiment with and without force feedback. In first case the system (SPIDAR-GL) was used while its haptic rendering was active then it was made inactive and users depended only on the visual feedback to manipulate objects in virtual environment. While the haptic rendering was active the users not only experienced the weight for all objects but also the resistive force the static obstacles (walls). In order to avoid any training transfer, these conditions were counterbalanced.

Initially all cubes were present in the middle of the two zones and near to the screen. The users were asked to randomly select a cube and place it all on a zone of their choice. Once all cubes were placed on the given zone, the participants were asked to place all the cubes on the other zone and finally again on the first one. To record the user's opinion about the system, they were asked to fill a questionnaire. The questions were the following: "the device used in the application is simple", "it is light, natural", "objection selection & positioning is easy, precise, and comfortable", "the system don't give tiredness", "the wires don't disturb visualization", "the task is easier with force feedback", "Each object has a weight" and "it is differentiable for two objects closed in mass (200g & 300g)". Each question has three options for response such as: agreed, neutral, not agreed.

The following graph in the figure 8 represents the results of the questionnaire.

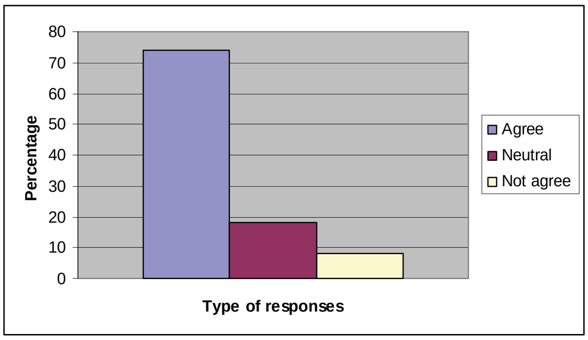


Figure 8. Results of the questionnaire.

7. Conclusion and future work

We presented an original integration of the SPIDAR-G (a string-based haptic device) in Virtools Dev 4.0 software environment. The overall system allows the users to interact with virtual objects or obstacles (walls) in a large-scale set-up. An experimental virtual environment was developed in order to carry out a usability study of this VR platform. Participants were asked to repetitively pick and place a total of nine cubes on the floor of a virtual room, using the SPIDAR-G haptic device. Results from questionnaire revealed a very good acceptance and feeling from the users. (73% of responses are positive, 18% of responses are neutral and only 8% of responses are negative). This shows that the SPIDAR in virtual environment has many advantages over other grounded force feedback devices in terms of usability and transparency. Such a device is therefore more likely to contribute to the field of enactive interfaces.

In order to investigate some other interesting characteristics of SPIDAR like speed, force and accuracy etc, we are going to carry out more precise experimentations in the near future. Similarly we are planning to deploy the SPIDAR for virtual/augmented reality supported teleoperations and collaborative teleoperations.

Acknowledgments

This work is supported by the *VarSCW* (Virtual and augmented reality Supported Collaborative Work) project. We wish to thank "*Le Conseil Général de l´ Essonne*", *CNRS* and the *IBISC laboratory* for funding this project.

References

[1] M. Sato "Development of String-based Force Display: SPIDAR**",** IEEE VR 2002 Conference.

[2] N.Tarrin, S.Coquillart; S.Hasegawa L.Bouguila and M.Sato "The string haptic workbench: a New Haptic Workbench Solution". Computer Graphics Forum, Volume 22 Issue 3, pp. 83-89, September 2003.

[3] G. Burdea "Haptic Feedback for Virtual Reality" Innovation Research, Vol. 2, No. 1, pp. 17-29, July 2000.

[4] **http://www.sensable.com**

[5] **http://www.immersion.com**

[6] L. Yanlin, J. Murayama, K. Akahane, S. Hasegawa and M. Sato" Development of new force feedback interface for two handed 6DOF Manipulation- SPIDAR-G&G System". ICAT 2003, Tokyo Japan, December 4-5 2003.

[7] P. Richard, D. Chamaret, F.-X. Inglese, P. Lucidarme, J.-L. Ferrier "Human-Scale Virtual Environment for Product Design: Effect of Sensory Substitution"**,** the international journal of Virtual Reality, Vol. 5, No. 2, pp. 37-44, 2006.

Postural coordination biofeedback during stroke rehabilitation

Deborah Varoqui* Benoît G. Bardy* Jérôme Froger† Jacques-Yvon Pélissier*†

(*) Motor Efficience and Deficience Laboratory, University Montpellier 1, France
(†) Service de rééducation et de réadaption neurologique, CHU Nîmes
E-mail : deborah.varoqui@univ-montp1.fr, benoit.bardy@univ-montp1.fr, jerome.froger@chu-nimes.fr, jacques-yvon.pelissier@wanadoo.fr

Abstract

Postural instability caused by hemiplegia is an important correlate of this pathology. Based on accumulated knowledge on stroke-related postural deficits, various rehabilitation programs have been developed. Training based on center-of-pressure visual biofeedback have been used to improve balance capacities. It has not been clearly proved however that these protocols can improve balance. COP kinematics contains interesting variables capturing postural stability but it does not allow to understand the origins of the observed deficits.

Here we proposed to study the multi-segmental postural system in hemiplegic population with concepts and tools issued from the pattern dynamics approach. Ankle-hip coordination patterns were investigated. Results showed that hemiplegic subjects do not produce the in-phase pattern and are less stable in the anti-phase pattern compared to healthy subjects.

From these results, we built a postural pattern (re-) learning system. The experimental set-up used, Virtual Posture, is a customized biofeedback system initially develop to study postural pattern formation in healthy participants. We present preliminary data with an hemiplegic population.

1. Introduction

Stroke is a neurological deficit of cerebrovascular cause which usually involves hemiplegia. Hemiplegia is characterized by paralysis of one vertical half of the body. In general, the defect in the brain is on the opposite side of the body. Patients present many different deficits such as abnormal muscular tone (e.g., rigidity or spasticity), sensory deficits, or weakening of motor control. All these deficits contribute to reduce balance. Impaired postural control is regarded as the major problem for hemiplegic persons because it increases the number of falls and limits autonomy [5].

Many studies have investigated postural deficits by analyzing the kinematics of the center of pressure (COP) in quiet stance. Results have shown that hemiplegic posture is characterized by an important postural asymmetry between the healthy side and the injured side (i.e., the healthy leg supports the major part of body weight) and a large increase of postural oscillations along the medio-lateral axis [10].

This difference between the two sides is also observed at the neuromuscular level. For example, it has been shown in perturbation studies that latency of muscle activation in the paretic leg increases and that timing between distal and proximal synergists is disturbed compared to the healthy leg [3].

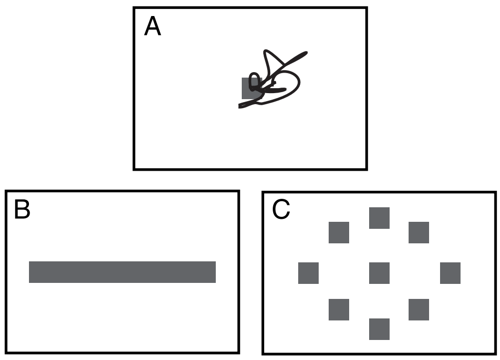


Figure 1: Illustration of different COP biofeedback tasks. Subjects are asked to (A) maintain COP (black line) in the grey target, (B) shift weight from left to right in the rectangle and (C) transfer weigh in order to point each target. Adapted from Nichols et al. (1997).

This data on postural deficits encourage adapted rehabilitation procedures. In the past decades, rehabilitation using various biofeedback techniques (bioFB) has developed and today complete classic physical therapies (e.g., EMG visual bioFB, auditory bioFB). For balance rehabilitation, visual bioFB from COP displacement is often used, in order to improve weight-bearing symme-

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

try and *dynamic stability*[1], two good indexes of stroke severity [9]. In this type of protocol, a computer screen gives real-time visual information about the patient's COP. Patients are asked to "control" their COP in order to do different tasks such as pointing in different directions (see Figure 1 for an illustration) [2,10,7]. In general, results show an improvement of weight-bearing symmetry [10]. But there is no long-term effect on balance and function independence when compared to classic physical therapy. For a review of bioFB training studies, see [5,7].

The COP is a good representation of the consequence of postural disturbance in stroke patients but it does not give direct access to the origins of posture deficits. Thus, it is possible that biofeedback-COP-based rehabilitation is not completely adapted to improve balance in hemiplegic patients.

2. Postural dynamics in hemiplegia

2.1 Introduction

Here we propose to study postural control using the coordination dynamics framework. In this approach, the multi-segmental postural system is considered as a dynamical system with multiple coordination modes. The relative phase ϕ_{rel} between angular motion of hips and ankles was analyzed for this purpose. In healthy population, two preferred modes of coordination were found to exist between these two joints [1]: an *in-phase* mode, with the two joints oscillating simultaneously in the same direction (ϕ_{rel} close to 0°), and an *anti-phase* mode, with the two joints moving in opposite directions (ϕ_{rel} close to 180°).

Emergence and stability of these preferred patterns depend three types of constraints: the environmental constraints, the intrinsic subject-related constraints and the task constraints [6]. For example, an experimental modification of the location of the center of mass (i.e., intrinsic constraint) affects postural dynamics. The higher the center of mass, the more the *anti-phase* pattern is adopted. All sensori-motor consequences of stroke can be viewed as intrinsic constraints that can modify coordination dynamics. For example, Rice and Newell (2004) showed the impact of hemiplegia on bimanual coordination and found that hemiplegic subjects have difficulties to achieve the *anti-phase* pattern. The *anti-phase* pattern corresponds to the less stable pattern for bimanual coordination in healthy subjects.

We report thereafter the results of two experiments in which we observed the changed in postural dynamics following stroke. We supposed that hemiplegia modifies postural dynamics. In experiment 1, we expected the less stable pattern (the *in-phase* pattern for the postural system) to vanish in the hemiplegic population. In experiment 2, we expected the coordination bioFB used to contribute to postural rehabilitation.

[1] *Dynamic stability* classically refers to the ability to move with a given posture without loss of balance; it is different from the concept of postural dynamics used later in this paper.

2.2 Method

Postural coordination dynamics was investigated in hemiplegic population (N = 12), and compared to healthy subjects (N = 12) placed in weight-bearing asymmetry with 70% of body weight on one leg. Participants were asked to perform two postural patterns – 0° and 180° – with the help of *Virtual Posture*, a customized biofeedback system initially developed to study the learning of new postural patterns [4]. The postural pattern to be achieved was visually represented by a Lissajous figure (i.e., ankle-hip position plane) displayed on a screen. The angular displacement of hips and ankles were recorded with two electrogoniometers (Biometrics Ltd.). These data were used to generate a real time visual bioFB displayed on the same Lissajous figure. Thus, participants could see the difference between the pattern required and the pattern produced.

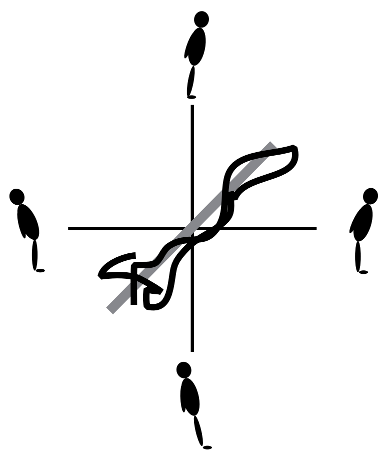


Figure 2: Pattern to produce (oblique line, corresponding to the 0°relative phase) and real time visual bioFB (grey line) represented on an ankle-hip position plane.

2.3 Results and Discussion

When 0° was the required pattern, hemiplegic subjects produced a coordination pattern close to 180° while healthy subjects presented a bidirectional distribution of relative phases with peaks centered around 0° and 180° (Figure 3). This bi-directionality appeared in the overall distribution because subjects achieved either 180° or 0° in a given trial. When 180° was the pattern required, the two groups successfully achieved the instructed pattern (For the hemiplegic group, mean ϕ_{rel} = 175.65°, $SD\phi_{rel}$ = 44.57° and for the healthy group, mean ϕ_{rel} = 172.20°, $SD\phi_{rel}$ = 15.22°).

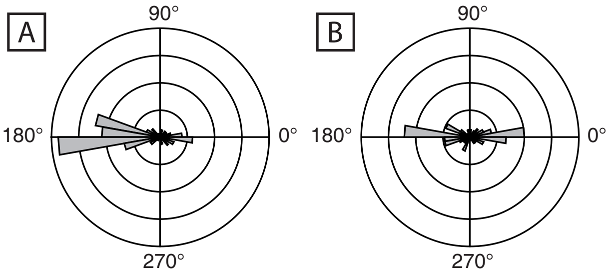


Figure 3: The occurrence of ϕ_{rel} produced by the hemiplegic/unload leg for the hemiplegic group (A) and the healthy group (B) when 0° pattern was required.

Standard deviation of ϕ_{rel} ($SD\phi_{rel}$) which stands for pattern stability, showed that the stability of coordination was smaller for the 0° task than for the 180° task. More precisely, healthy subjects were more stable than patients in the 180° task but less stable in the 0° task (Figure 4). In the latter condition, patients were less variable because they achieved the 180° pattern (known to be the most stable) instead of the required pattern. Moreover, coordination was more stable with the healthy side than with the injured side. For the healthy subjects, the imposed constraint clearly destabilized the *in-phase* pattern.

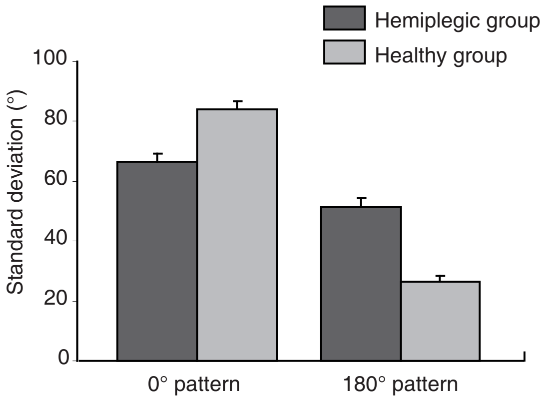


Figure 4: Stability and standard error of ϕ_{rel} for the hemiplegic/unload leg of the 2 groups.

In conclusion, hemiplegic subjects did not exhibit the 0° pattern and they were less stable than healthy subjects for the 180° pattern. Hemiplegia and its various consequences have modified postural dynamics. In addition, postural asymmetry, as introduced in the healthy subjects, was not a clear determiner of these modifications.

3. Re-learning patterns with biofeedback rehabilitation

From these results, we elaborated a re-learning protocol of preferred postural patterns while using *Virtual Posture*. This experimental set-up was used successfully in previous studies to learn new postural patterns in a healthy population [4]. We investigated 0° and 180° patterns because the first mode could not be produced and the second was less stable in stroke patients compared to a healthy population.

3.1 Method

Twenty-four hemiplegic patients engaged in post-stroke rehabilitation took part in this experiment. Participants were distributed in three groups: two experimental groups − group A and group B − performed a complete learning protocol with *Virtual Posture* and one control group − group C − realized only the pretest and the posttest (Table 1). The task using *Virtual Posture* was similar to the one used in the first study.
In the test periods (pretest and posttest), participants performed ten 60-s trials at the imposed pattern (i.e., 0° or 180°).with a bioFB coming from injured leg (iFB).
In the learning period (experimental groups only), subjects realized four sessions of practice with the imposed pattern. In each session, they were asked to produce twelve 60-s trials with different bioFB condition according to the group. The bioFB originated from the healthy side (hFB) for group A and from the injured side (iFB) for group B. The control group (group C) practiced a stand-up task during 15 min instead of learning patterns with *Virtual Posture*.
The protocol lasted 4 weeks (2 weeks per patterns) for all groups, during which subjects continued physical therapy. Balance, gait and functional independence were evaluated with different clinical tests at the beginning and at the end of the protocol to compare groups between them.

Table 1: Experimental protocol description for one pattern for each group (*VP*: use of *Virtual Posture* set-up for the training, hFB: bioFB from healthy leg, iFB: bioFB from injured leg).

	A	B	C
Pretest	10 trials **iFB** with *VP*		
Session 1	12 trials **hFB** with *VP*	12 trials **iFB** with *VP*	15 min Stand-up
Session 2	12 trials **hFB** with *VP*	12 trials **iFB** with *VP*	15 min Stand-up
Session 3	12 trials **hFB** with *VP*	12 trials **iFB** with *VP*	15 min Stand-up
Session 4	12 trials **hFB** with *VP*	12 trials **iFB** with *VP*	15 min Stand-up
Posttest	10 trials **iFB** with *VP*		

3.2 Results and Conclusion

Preliminary results on eleven subjects (5 of group A and 6 of group B) are somehow mixed, due to an important within – as well as between – subject variability. In the critical 0˚ condition, two different types of behaviour emerged. Six subjects learned the *in-phase* pattern at the end of the learning session, and five subjects still produced the *anti-phase* pattern (i.e., the pretest versus posttest comparison showed no difference for this group). These contrasting behaviours are illustrated in Figure 5 for two subjects.

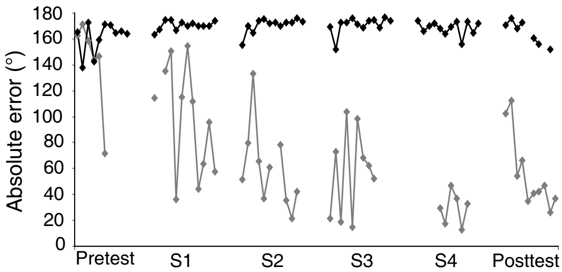


Figure 5: Mean absolute error per trial for two subjects learning 0˚, illustrating two distinct behaviours, a migration toward the requested 0˚ pattern (in grey) and a fixed 180˚ pattern (in black).

In conclusion, results show that some patients exhibited the requested *in-phase* pattern after four weeks of learning whereas others were unable to produce this pattern. The difference between the two groups in postural comfort at the end of the learning session, together with the correlation between group membership and type of hemiplegia (e.g., lesion type, affected side, etc.) are two open questions that will now be addressed.

References

[1] Bardy, B. G., Marin, L., Stoffregen, T. A., & Bootsma, R. J. (1999). Postural coordination modes considered as emergent phenomena. Journal of Experimental Psychology: Human Perception and Performance, 25, 1284-1301.

[2] De Haart, M. Geurt, A. C., Huidekoper, S. C., Fasotti, & L. Van limbeek, J. (2004). Recovery of standing balance in post-acute stroke patients: a rehabilitation cohort study. Archives of Physical Medicine and Rehabilitation, 85, 886-895.

[3] DiFabio (1987). Lower extremity antagonist muscle response following standing perturbation in subjects with cerebrovascular disease. Brain Research, 406, 43-51.

[4] Faugloire, E., Bardy, B. G., Merhi, O., & Stoffregen, T. A. (2005). Exploring coordination dynamics of the postural system with real-time visual feedback. Neuroscience Letters, 374, 136-141.

[5] Geurt, A. C., de Haart, M., van Nes, I. J., & Duysens, J. (2005). A review of standing balance recovery from stroke. Gait and Posture, 22(3), 267-281.

[6] Newell (1986). Constraints on the development of coordination. In M. G. Wade & H. T. A. Whiting (Eds.), Motor development in children: Aspects of coordination and Control (pp. 341-360). Dordrecht: Nijhoff.

[7] Nichols, D. S. (1997) Balance retraining after stroke using force plateform biofeedback. Physical Therapy, 77(5), 553-558.

[8] Rice, M. S., & Newell, K. M. (2004). Upper-extremity interlimb coupling in persons with left hemiplegia due to stroke. Archives of Physical Medicine and Rehabilitation, 85, 629-634.

[9] Sackley, C. M. (1990). The relationship between weight-bearing asymmetry after stroke, motor function and activities of daily living. Physiotherapy Theory and Practice, 6, 179-185.

[10] Shumway-Cook, Anson, & Haller (1988). Postural sway-biofeedback: its effect on reetablishing stance stability in hemiplegic patients. Archives of Physical Medicine and Rehabilitation, 69, 395-400.

This research was supported by grants from the *French Ministère de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche* and by the Enactive Interfaces European Network of excellence (IST contract ♯002114).

Evaluation of Directional Force Threshold through Psychophysics Experiments

Marco Vicentini Maria Carla De Maggio Debora Botturi Paolo Fiorini
Department of Computer Science
University of Verona, Italy
E-mail: {vicentini, demaggio, debora, fiorini}@metropolis.sci.univr.it

Abstract

Psychophysics experiments have been carried out in order to measure the differential thresholds of force perception applied to the hand-arm system, with the aim of better understanding the human perception capabilities. The experimental work consists in analyzing how force is feedback to the user using a 6 Degrees-of-Freedom haptic device. A Maximum-Likelihood adaptive procedure is used to identify the perceptual threshold of 13 subjects, recruited among the staff of our laboratory. Our findings indicate that the Just-Noticeable-Difference is generally higher than previously reported in the literature. Furthermore, we found evidences that the thresholds change not only along the stimuli continuum, but also among the different joints. These findings support our claim that human perceive forces differently along different directions, thus suggesting that perception can be enhanced by suitable signal processing.

1 Introduction

Haptic Interfaces were developed to simulate the action of touching virtual or remote objects endowed with suitable dynamics. This technology is applicable to many fields, such as teleoperation, robotic manipulations and computer simulation.
The force feedback provided by a haptic device is usually defined as the sensation of weight or resistance felt by a human operator [4]. It requires a device that produces a force to the operator such as the one of the interaction with a real object, allowing a person to feel the weight of virtual objects, or the resistance to motion they create.
Basic understanding of the biomechanical, sensorimotor, and cognitive abilities of the human haptic perception is critical for proper design specification of the hardware and software of haptic interfaces.
Several works are relevant to the quantitative measures of human factors that affect the design of force-reflecting haptic interfaces. One of the most common measures is related to the Just Noticeable Difference (JND), which is the minimal difference between two stimuli (F vs. $F + \Delta I$) that leads to a change in the human perceptual experience and is detectable by a human being.
The JND percentage value for pinching motions between finger and thumb was found to be around 7% of the reference force [9]; in a force matching experiment about the elbow flexor muscles, a JND ranging between 5% and 9% was observed [8]. The JND was found to be relatively constant over a range of different base force values between 2.5 and 10 N, and essentially independent from reference force and displacement [1].
In the following reviews [7, 10, 12], these results are presented as key in human force perception for the design of haptic device. However, these reviews do not adequately stress the fact that the results are focused on finger capabilities or that the experiments do not primarily investigate haptic environments, which have the native capability to produce accurate force stimuli.
The experimental methodology of recent studies [2, 11] often avoided considering the elementary force perception. For example, the perception of the force vectors angle was analyzed in Cartesian space as a whole, not with respect to the contribution of each Degree of Freedom (DoF) involved in generating the target force.
Our aim is to identify whether there are differences in terms of force perception among the directions and orientations in Cartesian space. This finding could let us distinguish the most sensible directions for the arm, allowing us to identify a suitable scaling matrix for force-feedback in haptic environments. Thus, the purpose of the experiments descrived in the following is to investigate the capability of the arm to discriminate a force along each Cartesian axis (x, y, and z) and the capability of the wrist to discriminate torques along each rotational axis ($roll$, $pitch$, and yaw).

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

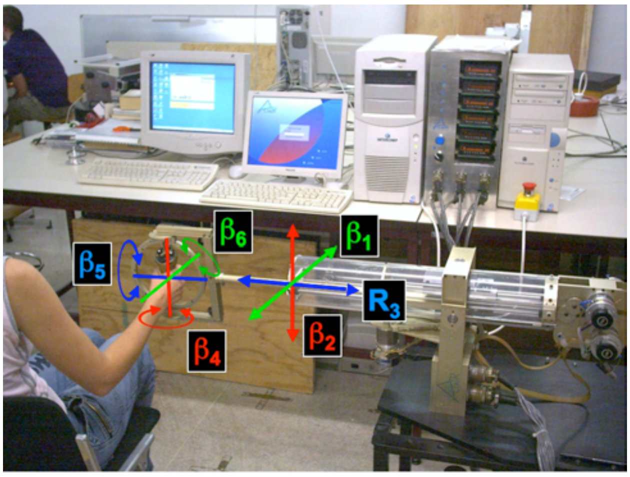


Figure 1: The experimental setup for the Force Reflecting Hand Controller haptic device. The Cartesian and rotational axes are accompanied by the names of the joint variables.

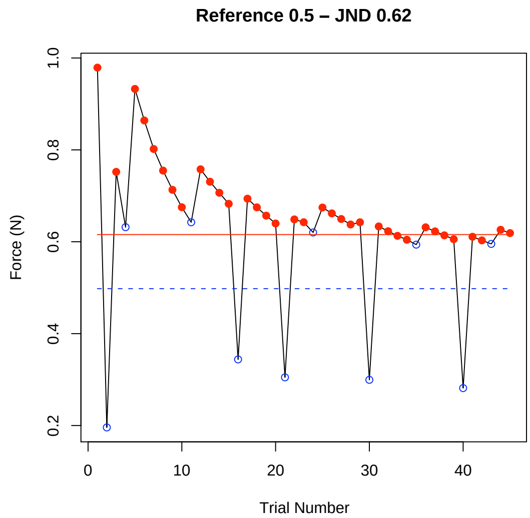


Figure 2: Adaptive track following the Green's adaptive procedure. Red filled points refer to positive responses; blue circles to negative ones. Red line refers to the estimated threshold; blue dashed line to the reference force.

2 Methods

2.1 Apparatus and Materials

NASA's JPL Force Reflecting Hand Controller (FRHC) was used in our task, because of its particular features [3, 5]. This device has 6 DoFs, consisting of a translational and five rotational joints (see figure 1). It slides and rotates around a fixed support attached to the floor, through a translational (R_3) and two rotational joints (β_1, β_2). It has a handgrip with three intersecting axes ($\beta_4, \beta_5, \beta_6$). It permits to operate with full dexterity in a cubic workspace of $30 \times 30 \times 30$ cm developing forces up to 10 N and torques up to 0.6 Nm.
Motion transmission is done by wires with pulleys of large curvatures that reduce friction and increase backdrivability. An idler mechanism translates on the opposite direction to R_3 to make the manipulator always balanced. The power unit is situated on the floor to avoid unbalancing the structure. The FRHC low friction combined with its low inertia increases the interface force fidelity and makes it the ideal instrument for the planned force perception experiments.

2.2 Participants

13 subjects (mean age of 26 years, age range from 19 to 36 years) were examined, almost all of them with no previous knowledge of the experiments. The participants were recruited among the staff of the Altair laboratory of the University of Verona (Italy) by word of mouth and did not receive any compensation for their participation. All the participants were right-handed, had a normal touch sense and used their dominant hand to perform the task.

2.3 Design and stimuli

The aim of the experimental design was to measure the capability of the human hand in terms of force perception. The Green's [6] maximum likelihood adaptive procedure of psychophysics measurement, in the 2 alternative forced-choice paradigm, was involved to measure the JND; this procedure promises highly efficient trial placement and threshold estimation.
During each trial, different controllable forces were applied in successive intervals to a single joint, generated by the force-feedback device. Subjects were given a choice of 2 alternatives, and they had to select the one containing the stimulus. The subjects know that exactly one alternative contains the stronger stimulus $F + \Delta I$, and that the other has the stimulus F.
For each joint, 10 reference forces F were identified among the physical stimulus domain. The comparison stimuli $F+\Delta I$, which would be presented to the subject, depended critically on the subject's responses. That is, while the reference force was constant during a trial, the comparison stimulus changed during the trial according to the subject's answer.
In this procedure, after stimulus presentation, a set of candidate psychometric functions was fitted to all the data collected up to that point, and the likelihood associated with each function was computed. The most likely function determined the comparison stimulus $F + \Delta I$ to be used on the next trial, followed by another updating of the candidate function probabilities. The final estimate of threshold was extracted from the most likely

psychometric function after some number of trials. In the example depicted in figure 2, the threshold level appears to stabilize after about 25 trials.
Each force stimulus was applied for 1,200 ms; the interval between stimuli was 300 ms. Subjects were instructed to keep the hand movement as small as possibile. After stimuli presentation, subjects were asked to judge which was the stronger one. The experimental session lasted about 30 minutes.

3 Results

The data collected were analyzed in order to estimate the perceptual threshold for the reference stimuli. For each joint, ANOVA was conducted in order to determine if there were significant differences among the perceptual thresholds given by individual differences. As shown in table 1, differences in the JND% among the subjects were observed for all but one translational and rotational axes; no significant differences were observed only along the axis β_1. Even so, attributable to several missing data, the statistical test conducted for the axes β_4 and β_5 should be cautiously considered.
In figure 3, the force JND% was plotted versus the reference force for one prototypical joint. We observed a non linear relationship between the reference stimuli and the JND%. Considering the lower reference stimuli, the force stimulus and the perceptual threshold appeared to be associated inversely: the lower the force or torque applied, the higher the threshold perceived.
In table 1, we have reported the median of the distribution of the estimated thresholds along each stimulus continuum. A difference between force and torque perception was observable for the low intensities: the perceptual thresholds of low torques were always greater than 40% and often over 100%. However, this result could also be due to the inertia of these joints. The perceptual thresholds for intense stimuli were higher than in the classical experiments: our findings provide some evidences that the perceptual threshold for intense forces or torques is around 15% - 20% of the reference force.
We have investigated whether the thresholds could be changed due to different joints. Repeated measures ANOVAs were performed (factors: n joints $\times$ 10 stimuli $\times$ 13 subjects) for all the n combinations of the joints. That is, we have considered all the combinations of joints among themselves, with elements $k = 2, \ldots, 6$, and for each combination the analysis of variance was conducted. The interactions among the factors joints $\times$ stimuli, and the factors themselves were always highly significant ($p < 0.001$). These results could lead to infer that different thresholds have to be considered along each joint.

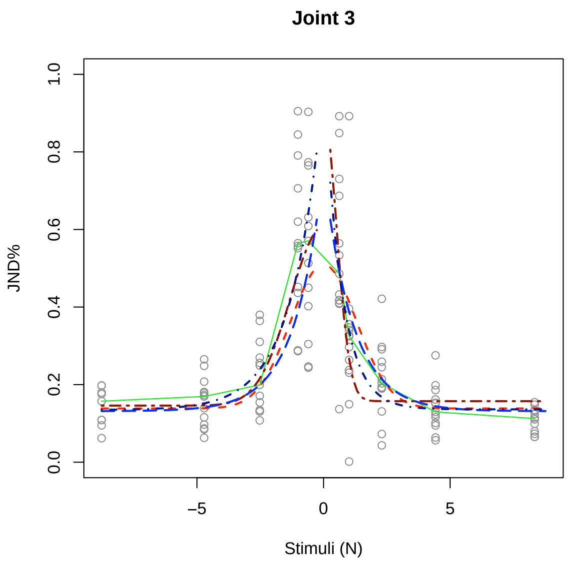


Figure 3: Force JND% versus reference force for joint R_3. Each point is the subjects' individual threshold. The green line maps the median values; for both the symmetric and asymmetric conditions, the blue curves (dot-dashed and long-dashed) refer to the exponential-type function F_1, while the red ones (dashed and two-dashed) to the bell-type function F_2.

4 Conclusions and Future Works

The goal of this work is to understand which is the force range better perceived by the human arm and wrist. With respect to the literature [9, 12], we obtained different results: mean threshold 15% - 20%. The force perception differs from joint to joint and also within the range of one joint, those results justify a deeper study on different perception and possible scaling.
Some future goals are the identification of new haptic device specifications and a new force scaling concept related to the master device at hand and to the type of application. For example, in surgery teleoperation low intensity signal are very common but not well perceived by the human, therefore a suitable signal processing is needed to enhance the perception and moreover the operation performance.

Acknowledgements

This work was partially founded by the AccuRobAs project under EU's 6th Framework Programme (contract IST-045201) and by the AIDA project (MIUR COFIN 2005-2007).
The authors would like to thank Dr. Gabriel Baud-Bovy,

Joints		Reference Stimuli										ANOVA Factor: Subjects
		s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	
β_1	N	-5.99	-3.28	-1.80	-0.87	-0.54	0.50	0.83	1.79	3.26	5.98	
	JND%	15.3	16.3	18.0	18.1	25.7	31.6	22.7	15.7	17.0	9.6	$F_{12,97} = 1.69$ $p = n.s.$
β_2	N	-7.21	-3.87	-2.06	-0.80	-0.51	0.41	1.07	2.29	4.07	7.35	
	JND%	11.62	12.56	19.29	36.93	25.33	29.58	32.66	24.87	13.25	12.16	$F_{12,101} = 2.46$ $p = 0.007$
R_3	N	-8.77	-4.71	-2.52	-1.01	-0.60	0.62	1.01	2.30	4.42	8.33	
	JND%	15.73	16.95	19.92	56.47	57.06	48.57	32.47	20.31	12.96	11.25	$F_{12,103} = 3.21$ $p < 0.001$
β_4	Nm $\times 10^{-2}$	-32.93	-16.87	-8.17	-4.50	-2.78	2.59	4.78	10.25	18.22	32.94	
	JND%	12.65	23.25	45.06	35.88	41.99	63.02	56.46	31.48	23.85	13.79	$F_{12,74} = 5.11$ $p < 0.001$
β_5	Nm $\times 10^{-2}$	-21.24	-12.02	-9.40	-6.50	-3.60	1.95	6.87	11.37	15.43	29.72	
	JND%	11.13	16.20	18.56	17.56	66.73	>100	46.14	31.61	28.19	/	$F_{12,85} = 2.16$ $p = 0.021$
β_6	Nm $\times 10^{-2}$	-41.22	-20.86	-15.07	-8.67	-3.50	4.84	11.86	18.86	25.20	47.48	
	JND%	12.11	19.30	21.16	24.02	> 100	46.04	30.86	27.48	20.49	12.21	$F_{12,91} = 3.56$ $p < 0.001$

Table 1: For each of the 10 reference forces identified along the force continuum of each joint the median value of the JND% distribution is reported. The ANOVA F-value refers to the significance of the mean difference among the subjects.

Mr. Stefano Galvan and Mr. Francesco Bovo for the experimental setup and their invaluable suggestions.

References

[1] S. Allin, Y. Matsuoka, and R. Klatzky, "Measuring just noticeable differences for haptic force feedback: Implication for rehabiliatation," in *Haptic Interfaces For Virtual Environments & Teleoperator Systems*. Los Alamitos, CA: IEEE Computer Society, 2002, pp. 299–303.

[2] F. Barbagli, K. Salisbury, C. Ho, C. Spence, and H. Z. Tan, "Haptic discrimination of force direction and the influence of visual information," *ACM Transactions on Applied Perception*, vol. 3, no. 2, pp. 125–135, 2006.

[3] A. Bejczy and K. Salisbury, "Kinematic coupling between operator and remote manipulator," *Advances in Computer Tecnology*, vol. 1, pp. 197–211, 1980.

[4] G. Burdea, *Force and touch feedback for Virtual Reality*. New York, NY: John Wiley & Sons, Inc., 1996.

[5] S. Galvan, D. Botturi, and P. Fiorini, "FPGA-based controller for haptic devices," in *IEEE Intl. Conf. on Intelligent Robots and Systems*. Los Alamitos, CA: IEEE Computer Society, 2006, pp. 971–976.

[6] D. M. Green, "A maximum-likelihood method for estimating thresholds in a yes-no task," *Journal of the Acoustical Society of America*, vol. 93, no. 4, pp. 2096–2105, 1993.

[7] V. Hayward, O. R. Astley, M. Cruz-Hernandez, and G. Robles-De-La-Torre, "Haptic interfaces and devices," *Sensor Review*, vol. 24, no. 1, pp. 16–29, 2004.

[8] L. A. Jones, "Perception of force and weight: Theory and research," *Psychological Bulletin*, vol. 100, no. 1, pp. 29–42, 1986.

[9] X. D. Pang, H. Z. Tan, and N. I. Durlach, "Manual discrimination of force using active finger motion," *Perception & Psychophisics*, vol. 49, no. 6, pp. 531–540, 1991.

[10] M. A. Srinivasan and C. Basdogan, "Haptics in virtual environments: taxonomy, research status and challenges," *Computer & Graphics*, vol. 21, no. 4, pp. 393–404, 1997.

[11] H. Z. Tan, F. Barbagli, K. Salisbury, C. Ho, and C. Spence, "Force-direction discrimination is not influenced by reference force direction," *Haptics-e*, vol. 4, no. 1, pp. 1–6, 2006.

[12] H. Z. Tan, X. D. Pang, and N. I. Durlach, "Manual resolution of length, force and compliance," in *Haptic Interfaces for Virtual Environment and Teleoperator Systems*, vol. 42. Anaheim, CA: ASME Dynamic Systems and Control Division, 1992, pp. 13–18.

An Architectural Platform for Audio-Haptic Simulation in Walking

Yon Visell*† Jeremy R. Cooperstock* Karmen Franinović†

(*) Centre for Intelligent Machines and CIRMMT
McGill University, Montréal, Canada
(†) University of Applied Sciences & Arts Zurich, Switzerland
E-mail: yon@cim.mcgill.ca, jer@cim.mcgill.ca, karmen@zero-th.org

Abstract

This paper presents work on the EcoTile*, a new interactive platform for the simulation of auditory and haptic experiences of everyday materials in walking. Walking is noteworthy as an enactive activity that is continuously negotiated in direct contact with its surroundings, in a way that is strongly dependent on the mutual morphology and material properties of both. As a situated activity, it is charged with many sources of meaning, and with potential for intervention at responsive and informational levels, although to date comparatively few devices have taken advantage of the unique opportunities it provides. The main design considerations behind the EcoTile are described, followed by the system organization and components, and initial experiences with the prototype.*

1. Introduction and context

Walking is an activity common to most humans, and to an impressive range of activities ranging from the functional to the recreational. It also carries significance for our health and well-being, and the study of walking and its disorders occupy a prominent role in the movement sciences. Walking has been prominently read by cultural theorists and philosophers including Michel de Certeau and Walter Benjamin as intimately tied to both the experience of space and the manner in which it is cognitively constructed through movement [1, 4].

It is surprising that more attention has not been devoted to the engineering of interfaces that explicitly make use of walking as an enactive activity. Recent years have seen a growth in invention in the area of mobile information devices that are adapted to movement by foot. For example, the Nike+ running shoe and music player system provides a sensor interface that lends an added informational dimension to the activities of runners. However, comparatively few devices have sought to take advantage of the unique qualities that accompany such modes of movement, whether at the sensorimotor level or the semantic one.

Our research is directed toward the design of floor-based interfaces whose sensing and actuating capabilities furnish dynamic control over the ecological information that is available in walking, interactively shaping the perceived auditory and haptic material qualities of the ground. A key source of motivation for this research is the possibility of merging new technologies for generating virtual experiences of material with the intrinsically tangible, physical, and spatial experience of walking. An advantage of foot based interaction lies in the possibility to engage intimately and individually with the senses and activities of people in public and semi-public spaces in ways that are difficult with conventional displays. For reasons such as these, we have elected to focus on the development of an architectural interface for enaction in walking, as one that can be utilized for the design of diverse spaces, without the kinds of restriction that would result from requiring users to don special equipment to experience it.

2. Related Work

Informative ground materials for walkers have long played a role in the design of urban environments. Passive haptic indicators are commonly used to mark important locations including stairways, crosswalks or subway platforms, and outdoor paths are designed to be readily distinguished by their material properties. However, the possibility of conveying these kinds of information through active devices embedded in real spaces in which people walk remains basically unknown. Devices that have been developed for interactive simulation in walking have been typically confined to laboratories or other closed environments.

Many researchers have investigated aspects of walking in the context of location-based interaction with a personal mobile information device. A few have addressed the role of computing in sensorimotor level in-

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

teraction, such as the studies by Crossan et al on interactions between gait phase during walking and simultaneous target selection on a mobile device [3]. Further, there is extensive prior research on medical applications of walking based interfaces, particularly in the diagnosis, care, and rehabilitation of disorders affecting gait (eg. [7]).

Several research groups have developed instrumented floors for the sensing of walking, dancing, or running movements. For an overview, see chapter 2 of the book by Wanderley and Miranda [6]. The PhOLIEMat developed by Cook [2] is somewhat unique as a floor-based device for the control of synthesized walking sounds via the feet (evoking the work of the Foley artist in film). To our knowledge no research has yet addressed the design of a floor for interactive simulation of surface material properties in walking, whether through the auditory or haptic channel.

As noted below, this work draws directly on the work of Fontana and Bresin on the physically based modeling of crumpling sounds [5].

3. Device Design and Characterization

A floor component called the *EcoTile* has been created with the aim of supplying interactive audio-haptic simulations of ecological ground materials in walking, based on a model for the simulation of the physical phenomenon of crumpling. The prototype device was designed with the aim of exploring the capability of an actuated but otherwise rigid interactive platform to interactively simulate the experience of walking on various materials. This aim was partly motivated by the notion that the auditory information alone is capable of conveying considerable information about the properties of surfaces that are walked upon. As a result, an even modestly successful haptic stimulus that is highly correlated with the auditory signal may be sufficient for a convincing percept.

In the prototype, the user walks on or otherwise interacts via the feet with the tile, which interactively delivers a signal designed to mimic the sensation provided by a given surface type. The stimulus consists of a vibration transmitted to the foot and simultaneously an acoustic signal transmitted to the ear, but originating near the platform. For this prototype, a snowy surface was selected as phenomenologically interesting and as a material that might be readily identified by the user through interaction[1]. The device is pictured in the images of Figure 3.

The mechanical component is a physical tile of dimensions 34 cm by 34 cm, together with a linear mech-

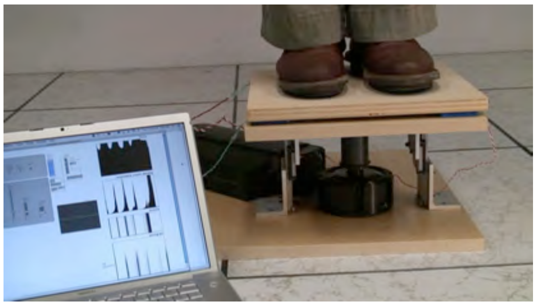


Figure 1: The interactive floor tile with user walking on it, next to the software component running on a laptop computer. The display of the latter shows the force, event density, and a sonogram of the resulting audio/haptic signal.

anism that allows it to be driven in a purely vertical direction. This platform is haptically driven with a motion control actuator from D-Box technologies. Sensing of the force applied by the user to the tile is accomplished with four Interlink force sensing resistors (FSRs) connected to a small microcontroller platform. The physically based simulation of the material is run in real time on a personal computer, which synthesizes the sound and haptic signals in response to the force on the tile, in the manner described below. The software environment used for prototyping is Max/MSP by Cycling'74. The sound is reproduced by a powered loudspeaker located at the base of the tile.

In the simulation of aggregates like snow, as implemented in the prototype device, the system can be controlled in open loop fashion. The required haptic force feedback through the device is much smaller than the force applied by the user (very roughly the user's weight), so the feedback from actuator to sensors is virtually negligible. Also, effects such as the displacement of the material underfoot (as in the compression of snow), which would likely require a closed loop control strategy, are ignored.

3.1 Physically Based Auditory and Haptic Stimulus Generation

Interactive auditory and haptic stimuli are generated continuously in response to the users footsteps. These are synthesized in real time by means of a physically based signal simulation of the phenomenon of crumpling, in a way that is driven by the force data from the footstep. An overview of the approach is presented in Figure 3.1.

The approach that was adopted utilizes prior work on the sound synthesis of such phenomena by Fontana et al [5], who have described the algorithm in significant detail. The model conceives crumpling pheomena to be

[1] Audiovisual documentation can be viewed at http://cim.mcgill.ca/ỹon/HS-Video – naturally, it does not capture the level of haptic interactivity.

composed of microscopic crumpling events, which are modeled as physical impacts between colliding objects, building on an earlier impact model by the same group [8]. Its main features are a stochastic model (Poisson process) governing the rate of production of such events, an energetic model positing the way the available energy is consumed, and a third component that addresses the change in the frequency characteristics of crumple events as crumpling unfolds. Parameters of the synthesis model may be tuned by hand to approximate various common or unusual crumpling artifacts or, as in the case of the current work, common or unusual aggregate ground surfaces.

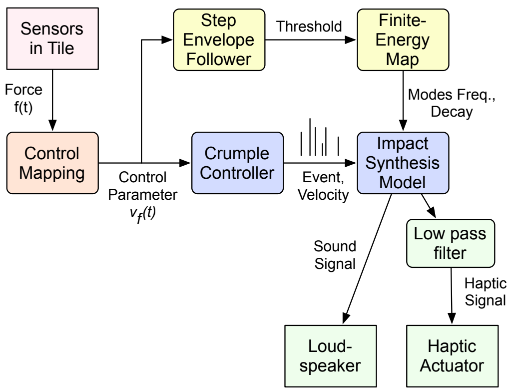


Figure 2: Diagram illustrating the crumple synthesis instrument, with force mapping, simulated crumpling model, and finite energy map.

In the prototype platform, both the auditory and haptic stimuli are derived from the same instance of the crumpling synthesis model. Due to the bandwidth limitation of the haptic actuator, the haptic signal was subjected to a low pass filter removing frequencies above 100Hz.

3.1.1 Force Control of the Crumpling Model

In the present application, a continuous control scheme providing a map from sensed foot pressure to sound synthesis parameters was needed, in order to maintain the highest possible level of interactivity. A parameter v_f used to control the amount of crumpling is obtained from the net force $f(t)$ exerted by the foot on the tile via:

$$v_f = \begin{cases} |df/dt| & \text{if} \quad df/dt < 0 \\ 0 & \text{else} \end{cases} \tag{1}$$

This choice represents a practical simplification in which crumpling occurs during walking only as weight is being transferred from the foot onto the tile, ignoring sensations produced as weight merely shifts from one part of the sole to another. The variation in force on the tile generated by such shifts does nonetheless contribute a qualitatively appropriate crumpling, as illustrated in video clips at the site mentioned above.

To implement this continuous control mapping, the Poisson process governing the generation of crumpling events was implemented such that the force v_f exerted on the tile at each instant determines, in appropriate units, the probability $P(\tau)$ that the time interval between crumpling events takes the value τ:

$$P(\tau) = v_f \, e^{-v_f \tau} \tag{2}$$

When the interval τ is small, the event probability is approximately v_f. The force f is polled from the device at intervals of $\tau \approx$ 10ms, and the distribution $P(\tau) = v_f$ is sampled to determine whether a crumpling impact is generated in the respective interval.

The energetic decay and frequency evolution of the crumpling model utilize the finite energy model Fontana et al, with the addition of an automated reset mechanism. In our system, the model state is reset at the beginning of each step by a simple footstep onset detector, consisting of a two-state envelope follower with a short attack and with a release time of approximately one second.

4. Qualitative Assessment and Future Work

In the prototype, auditory and haptic stimuli are derived from the same instance of the crumpling model. Due to actuator limitations, the haptic signal lacks frequency components above 100 Hz, whereas the human haptic channel is sensitive to those up to 1000s of Hz. Based on feedback from users, the auditory signal with the limited haptic signal appeared sufficient to convey the impression of snow. Ten users were asked to explore interaction with the device by stepping onto it, in stride or in isolation, with one or both feet. Several expressed surprise at the level of veridicality of the experience of the virtual snow. We felt this could be attributed to the synchronization between feedback channels, to the realism of the sound signal, and to the unexpected nature of interacting with virtual snow in a laboratory.

It was informally observed that the fusion of haptic and auditory modalities into a single percept could be completely disrupted by moving the loudspeaker farther than about 30 cm from the base of the tile, and then completely restored by placing the loudspeaker next to the base of the tile again.

The D-Box motion control actuator used in this prototype, while offering advantages including excellent response at low driving frequencies, has some drawbacks for this application. In particular, the control interface introduces a large latency in response that make it prohibitive to operate in situations that require a fast or closed-loop response, as in the case of hard surfaces.

A 100-tile floor 3m by 3m in size is currently under development. A new, modular tile is being designed to utilize a lower-cost and wider bandwidth, commercially available vibrotactile haptic actuator. The industrial design of the tile will be improved to provide a more solid and integrated platform that is protective of its key components, in order to permit testing with significant numbers of users in diverse settings.

Experiments are in progress with colleagues in psychology, utilizing real ground materials to assess the multisensory factors that contribute to material properties perception in walking. A data collection effort to document the physical signals present in walking on various materials, including applied force, haptic and auditory feedback. The aim is to make it possible to identify physically salient synthesis model parameter settings and control mappings from foot force profile to auditory and haptic response in the EcoTile simulations.

4.1 Future Application Scenarios

Unique applications of the EcoTile are anticipated in the architectural design of public and semi-public spaces, toward enabling the dynamic shaping of the aesthetic experiences of touch and material perception during walking, for functional purposes or with more creative aims in mind. A central question is tied to the possibility of invisibly organizing space and materiality in relation to the experience of walking or otherwise moving on foot within or through it. Functional applications are foreseen, including the reinforcement of sensory qualities for architectural spaces, and the creation of navigational aids or suggestions in the form of active, ecologically-based markers. Likewise, the tile may enable the creation of identity for a space or an organization (for example, the sensation of snow covered ground outside an outdoor store).

Other modes of creative production may be expressed through ground materials that morph in relation to the passage of people – emphasizing or de-emphasizing areas of passage, growing islands in places of congregation, or traces where people have passed. In short, such an interaction method may be thought of as charged with the possibility of lending these surfaces the range of affordances of the material world, but in ways that can be dynamically and interactively shaped. Dedicated structures such as virtual labyrinths might be constructed, as in a medieval church, with invisible paths representing meditation or contemplation, or locations might be linked over a network to other locations, lending one place the presence of those passing through another distant one. Further benefits of the technology being developed could lie with applications that can take advantage of an increased level of perceptual immersion, as in simulation training for rescue or other operations, physical or cognitive therapy, and interactive entertainment for theme parks, interactive science museums, or other play areas.

Acknowledgements

We thank the VIPS lab at U. Verona [5], who developed the sound models adapted for this work, and Don Pavlasek at McGill, who enabled the mechanical design. Bruno Giordano and Stephen McAdams have contributed to the expansion and perceptual foundation of this research.

This work is supported by the Centre for Interdisciplinary Research in Music, Media and Technology at McGill University, and the Hexagram Institute in Montreal. Additional support is provided by the CLOSED project, part of the European Commission's Sixth Framework "Measuring the Impossible" initiative: http://closed.ircam.fr

References

[1] W. Benjamin. *The Arcades Project*. Harvard UP, Cambridge, 1999.

[2] P. R. Cook. Modeling Bill's gait: Analysis and parametric synthesis of walking sounds. In *Proc. Audio Engineering Society 22 Conference on Virtual, Synthetic and Entertainment Audio*, Espoo, Finland, July 2002. AES.

[3] A. Crossan, R. Murray-Smith, S. Brewster, J. Kelly, and B. Musizza. Gait phase effects in mobile interaction. In *CHI '05 extended abstracts*, 2005.

[4] M. de Certeau. *The Practice of Everyday Life*. U. California Press, Berkeley, 1984.

[5] F. Fontana and R. Bresin. Physics-based sound synthesis and control: crushing, walking and running by crumpling sounds. In *Proc. of the XIV Colloquium of Musical Informatics (XIV CIM 2003)*, 2003.

[6] E. R. Miranda and M. M. Wanderley. *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. A-R Editions, Middleton, WI, 2006.

[7] J. A. Paradiso, S. J. Morris, A. Y. Benbasat, and E. Asmussen. Interactive therapy with instrumented footwear. In *CHI '04 extended abstracts*, pages 1341–1343, 2004.

[8] Matthias Rath and Davide Rocchesso. Continuous sonic feedback from a rolling ball. *IEEE MultiMedia*, 12(2):60–69, 2005.

Zooming experience in the haptic modality

Mounia Ziat*† Olivier Gapenne* Charles Lenay* John Stewart*
(*)University of Technology of Compiegne, France
(†)Center of Intelligent Machines, McGill University, Canada
E-mail: {mouni.ziat, olivier.gapenne, charles.lenay, john.stewart}@utc.fr

Abstract

The objective of this work concerns the design and the implementation of a zoomable interface implying the haptic modality. The initial postulate is that the zoom experience is not a natural, a direct experience, but supposes instrumentation and learning. In other words, the zoom experience is built by the appropriation of a technical substitution which makes it possible to modify the properties of the space-time flow; these properties which bind the subject to his (real or virtual) world are relational. To conceive this new interface, directly inspired from technologies known as of sensory substitution, we carried out a set of experiments allowing to define and to qualify the technical conditions and of use which favour the emergence of a perceptive experience of the zoom type. More generally, it concerns the proposition of more intuitive or immediate modes of instrumental interaction engaging explicitly the body in action.

1. Introduction

By using sensory substitution theories [1, 2, 3, 12], this work resumes the principle results allowing: i) to define a "haptic zoom" which encourages an active perceptive constitution and ii) to determine factors which are susceptible to lead to this perception. The choice of sensory substitution devices is crucial insomuch as they allow studying percept and space constitution through the subject activity and they underline the essential role of the technique which contributes tightly to this perception. Thus, a perception through a technical device is possible but it is constrained by the mediation of this device during the subject activity. The zoom, as a subject of study, is interesting and pertinent in the perception field since a zoomable perception exists only through the technical device. In other words, it supposes a mediation which is actively manipulated. And if this zoomable perception seems nowadays "direct" and "natural", one always forgets that it was the subject of a collective construction (from the 16th century) and a technogenesis. The zoom is a constituted modern perceptive experience, leading the subjects to forget the device creating this experience.

In one hand, while using a visual 2D½ or 3D zoom (prospective information on the depth is present), our body is engaged in this zoomable landscape and we have the feeling to advance, to move or to navigate in this landscape. In the other hand, while using a visual 2D zoom (zooming on a fixed image), the device is of course forgotten, but we have the feeling of being motionless relatively to the zoomable object. We are static but the object is not, it is approaching or moving away. In this case, the prospective information [8] informs on the magnification but not on the depth.

From these observations, we were interested by a numerical zoom implying the haptic modality. Our first hypothesis is that a zoom, which is generally visual, can be transposed to other perceptual modalities. We have tried to design a "haptic zoom" and to define its conditions of appropriation and use.

The designed zoom can be virtual or real[1] and the both present disadvantages and advantages which will be developed in the following sections.

2. Haptic zoom

In this study, the zoom is either an image expansion (real zoom) or an expansion of the virtual sensor (see below). Indeed, we can consider that the zoom function is a relative ratio between the perceived image and the sensor. While zooming, one modifies this ratio either by increasing the image size inside the capture

[1] The two zooms are operated in the numerical (virtual) space. However by "real zoom" we mean a real expansion (change of size) of the objects on the screen which is perceptible in term of distances. In the case of the "virtual zoom", the object size remains fixed because the size change is carried out on the virtual sensor. This last is considered as virtual because the perception of distances is possible only under certain conditions.

windows or, by decreasing the windows size and keeping the image size fixed (see figure 1).

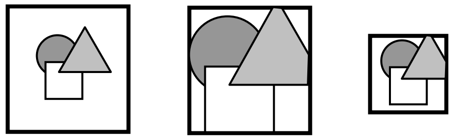


Figure 1. a) Initial image, b) real zoom, c) virtual zoom.

3. Experimental device

Tactos [10, 11] is a platform which allows the exploration of digital 2D shapes on a computer screen using tactile stimulations of the index finger. It includes three parts: a computer, a graphics tablet with stylus, and tactile stimulators (see Fig. 2). The stimulators are two electronic Braille cells, each including eight tactile pins. They are connected virtually to a sensor able to distinguish figures from the background on the computer screen. In other words, when the virtual sensor is on the outline of the figure a signal is transmitted to the stimulators and the corresponding pin is raised. The idea is to move the stylus on the graphical tablet so that a figure on the computer screen can be explored and recognized even though the user is blindfolded.

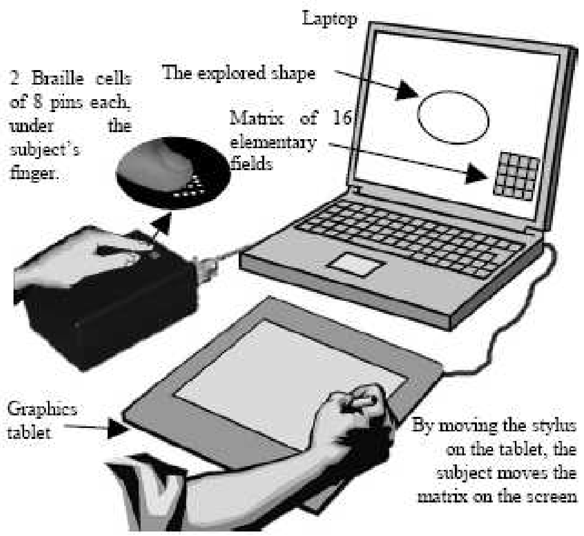


Figure 2. Experimental device: Tactos

4. Main results

Different experimental situations have been led in order to define factors which encourage a zoomable perception through the haptic modality.

4.1. The smallest perceptible

The zoom is a variation on the axis of scales [5] which is infinite but which is limited by the perceptual capacity of the human being. At the haptic level, the zoom experience is submitted to the same conditions: human movements which allow this perception are limited by physiological characteristics.

The first result of the first experiment [16] shows that the perceptual limit of subjects through a sensory substitution device is between 0.03 and 0.05 mm. these values correspond to the smallest displacements that a human being can produced. This result indicates that the motor resolution of 0.069mm [7] can be risen above. In our case, the absence of the visual modality can explain this rise.

4.2. Space perception

The haptic zoom implicates two spaces: it can be corporeal (real), i.e. the scale variation in on the shape (zoom on image); or it can be virtual (numerical), i.e. the scale variation in on the matrix (zoom on sensor). In the first case, the zoomable experience is happened in the corporeal space of the subject where their exploratory movements are adjusted according to the scale change of shapes. In the second case, the zoomable experience is happened in the virtual space, i.e. the screen space which is the space of interaction of objects and matrix which is displaced by the subjects; in this case, object sizes change relatively to the other virtual objects but stay unchanged in the corporeal space [17].

4.3. Speed reduction

The reduction of the speed displacement during the exploration activity is an important factor in the case of the virtual zoom. Indeed, for this last, any scaling is operated in the virtual space and the object size does not change in corporeal space.

To make this zoom “palpable”, the subjects reduce spontaneously their exploration speed at each zoom-in (reduction of the matrix size) and thus live the scaling experience. This speed reduction was already noticed during a precedent experiment [16] where the subjects changed their speed from 0.12 m/ms to 0.05 m/ms in order to be able to perceive the smallest sizes of letters.

This point was the subject of a more detailed study [19], and allows to release two profiles of subjects: i) the subjects which perceive a scaling, in this virtual case, because they succeeded in reducing sufficiently their movements according to the reduction of the matrix size and ii) the subjects which did not live the scaling experience because their speed remained quasi constant even higher according to the reduction of the matrix size.

4.4. Nature of task

The performances of the subjects vary according to the chosen task. Indeed, during different experiments, subjects were to carry out tasks of recognition of two types: a task limiting only to the recognition [16, 19, 20, 21] and a task requiring an additional estimation of distance [15, 17, 18].

The Table 1 summarizes the performances obtained according to the nature of task (OD: without distance estimation, WD: with distance estimation).

As shown on Table 1, performances in task OD are largely higher than the performances requiring distance estimation. Moreover, if we notice that the use of a virtual zoom is as efficient as a real zoom during a OD recognition task (1, 2a, 4a and 4b) it is not the same for a WD recognition task (3a and 3b); a zoom on image seems much more judicious for this type of task. This is due to the ambiguity between the corporeal space and the virtual space. Moreover, the performances for a real zoom are definitely higher because the change is palpable in the corporeal space (3c) and allows more precise distance estimation.

Table 1. Subject performances according to the nature task (WD and OD), (R for real zoom and V for virtual zoom)

Exp.	Task nature		Conseq.	Perfor.
	OD	WD		
1 [16]	R & V		↑	62 %
2a [19]	R & V		↑	70 %
3a [15]		V	↓	33 %
3b [17]		V	↓	33 %
3c [18]		R	↑	67 %
4a [21]	V		↑	80 %
4b [20]	R		↑	80 %

4.5. Proprioceptive experience

The proprioceptive experience is in close relationship to the range of selected scale. Indeed, the range of scales plays an essential part in the zoom perception. Subjects do not perceive the scaling for objects whose size is lower than 2.4mm because they have the impression to cross the same distance [19]. These small movements request hand and wrist movements. Contrary, for higher object sizes, scaling are perceived proprioceptively because the movements also request the front-arm.

4.6. Zoom-in, zoom-out and intermediate levels

When the subjects handle a haptic continuous zoom [20, 21] as on visual zoomable interfaces [4, 9, 13], they do not have any difficulty in make the distinction between a zoom-in and a zoom-out. In the same manner to a visual zoom, the zoom-in gives access to the detail and to position with precision on the object whereas the zoom-out, little used, is it used only to give a global view of the scene and allow them to centre the cursor on the objects present in the space of perception.

Another result showed that the maximum number of handled levels seems to be of 25. Indeed, subjects do not handle more than 25 levels and this, even if they have the possibility of handling some more (100 and 1000 levels) [20, 21].

5. Conclusion

In one hand, these experiments helped us to better understanding the constitution of a zoomable perception which corresponds to an alternation of expansion and depth when it is visual and by either a real displacement on a surface or the displacement of an object relatively to another when it is haptic. In the other hand, they reinforced us in the idea that a prosthetic perception is submitted to the same laws than the others perceptual modalities. Perception by means of prosthesis is a new perceptive experience which can be subscribed in our collective experience (at the enactive meaning, see [14]) by its learning as well as a “natural” perceptive modality.

The study of the zoom, as a perceptive experience mediated by the technique, enabled us to release that, when this last contains prospective information on the depth, it “affords” (In the Gibsonian meaning, see [6]) the locomotion, and that the subject, if he/she does not move, is engaged with his/her body in this virtual landscape. It is one of the essential reasons which make that the subject is in a perception of type “I approach / I move away from the object” and not of the type “I increase / I reduce the object”. This last case is the case of a zoom which do not contain any prospective information on the depth such as for example a 2D zoom.

We also highlighted that, if the zoom is a perceptive experience often conceived as visual, it is possible to constitute this perception with another perceptive modality (haptic in our case) whereas the tactile is not favourable to a depth perception in the real world since this sense requests enormously the contact with the object. In other words, the depth experience can be the subject of a true substitution within the framework of the interaction of the subjects with virtual objects.

Finally, we noticed that, even if a virtual haptic zoom is theoretically equivalent to a real haptic zoom, the two situations prove to be different and even contradictory in certain situations. Equivalence is true only when the subject is confronted to task recognition but disappears when the task integrates a distance estimation.

References

[1] Auvray, M. Immersion et perception spatiale. L'exemple des dispositifs de substitution sensorielle Psychologie cognitive, Ecole des Hautes Etudes en Sciences Sociales, Paris, 2004, 302.

[2] Bach-y-Rita, P. Brain Mechanisms in Sensory Substitution. Academic Press, New York, 1972.

[3] Bach-y-Rita, P. Sensory substitution and qualia. in Proust, J. ed. Perception et Intermodalité, PUF, Paris, 1996, 81-100.

[4] Bederson, B.B. and Hollan, J.D., Pad++: a zoomable graphical interface system. in Conference companion on Human factors in computing systems, (Denver, Colorado, USA, 1995), ACM Press, 23-24.

[5] Furnas, G. and Bederson, B.B., Space-scale diagrams: Understanding multiscale interfaces. in Proceedings of the ACM Conference on Human-Computer Interaction: CHI'95, (1995), ACM Press, 234-241.

[6] Gibson, J.J. The Ecological Approach to Visual Perception. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1979.

[7] Gosselin, F., Optimisation des interfaces haptiques : problèmes, méthodes, applications. in 17ème Congrès Français de Mécanique, (Troyes, France, 2005).

[8] Guiard, Y. and Beaudouin-Lafon, M. Espace et échelle du mouvement : la loi de Fitts dans le monde physique et dans les mondes électroniques. in Thinus-Blanc, C. and Bullier, J. eds. Agir dans l'espace, Maison des Sciences de l'Homme., 2005.

[9] Guiard, Y., Beaudouin-Lafon, M. and Mottet, D., Navigation as Multiscale Pointing: Extending Fitts' Model to Very High Precision Tasks. in Proceedings of the ACM Conference on Human-Computer Interaction (CHI'99), (Pittsburgh, PA, USA, 1999), ACM Press, 450-457.

[10] Hanneton, S., Lenay, C., Gapenne, O., Vermandel, S. and Marque, C., Dynamique de la reconnaissance de caractères via une interface haptique. in Actes du colloque ARC'98, (Saint Denis, Paris, 1998), Kayser, D. Nguyen-Xuan, A. Holley, A. (Eds), 343-347.

[11] Lenay, C., Gapenne, O. and Hanneton, S. La suppléance sensorielle : Limites et perspectives. in Hatwell, Y., Streri, A. and Gentaz, E. eds. Toucher pour Connaître : Psychologie Cognitive de la Perception Tactile Manuelle, PUF, Paris, 2000, 287-306.

[12] O'Regan, K. and Noë, A. What it is like to see: A sensorimotor theory of visual experience. Synthese, 129 (1), 2001. 79-103.

[13] Pook, S., Lecolinet, E., Vaysseix, G. and Barillot, E., Context and interaction in zoomable user interfaces. in Proocedings of the 5th International Working Conference on Advanced Visual Interfaces (AVI'00), (Palermo, Italy, 2000), ACM Press, 227-231.

[14] Varela, F., Thompson, E. and Rosch, E. L'inscription corporelle de l'esprit : sciences cognitives et expérience humaine. Seuil, Paris, 1993.

[15] Ziat M. Conception et implémentation d'une fonction zoom haptique sur PDAs. Expérimentations et usages, Université de Technologie de Compiegne, Compiegne, November 29th, 2006, 237.

[16] Ziat M., Gapenne O., Lenay C., Stewart J. Acuité perceptive via une interface pseudo-haptique, 16ème Conférence Francophone sur l'Interaction Homme-Machine (IHM'04), Namur, Belgique (30 août 30 - 3 septembre, 2004), pp. 263-266.

[17] Ziat M., Gapenne O., Lenay C. and Stewart J. Perception in numerical and corporeal spaces in a 2D haptic virtual world, 2nd International Conference on Enactive Interfaces (ENACTIVE'05), Genova, Italy (November 17-18, 2005). 10p.

[18] Ziat M., Gapenne O., Rouze M.-O. and Delwarde A. Recognition of different scales by using a haptic sensory substitution device, Proceedings of Eurohaptics'06, Paris, France (July 3-6, 2006).

[19] Ziat M., Gapenne O., Stewart J., Lenay C. Haptic recognition of forms at different scale: A comparison of two methods of interaction, Interacting with Computers, Elsevier (CV).

[20] Ziat M., Gapenne O., Stewart J., Lenay C. and Bausse J. Design of a Haptic Zoom: levels and steps, to appear in the proceedings of Worldhaptics'07, Tsukuba, Japan (March 22-24, 2007).

[21] Ziat M., Gapenne O., Stewart J., Lenay C IHM2007

Enaction_*in*_Arts Papers

KEYNOTE
-
Aesthetic Touch

Rosalyn Driscoll
Independent Scholar/Artist, USA
E-mail; rd@rosalyndriscoll.com

Abstract

The sense of touch provides information about people, objects and the environment—the world around us. Less known but equally important, touch gives us information about our internal environment—the world within us. Access to both subjective and objective realities is also integral to the experience of art. This capacity in both art and touch to reveal both subjective and objective dimensions makes touch deeply compatible with the experience of art. In making and exhibiting haptic sculpture, I have learned that this profoundly different way of knowing reveals new insights into art, perception and ourselves. Haptic perception is slow, linear, intimate, embodied, and concrete. It is reciprocal: as we touch we are also touched. This reciprocity creates a sense of mutuality and dialogue that deepens our experience of the world and of ourselves.

My response as an artist to the growing need for a more embodied way of knowing has been to make artwork that can be touched. To touch a sculpture is such an obvious, simple, natural act, something the artist does all the time and we all do when we can. Yet this simple act is actually quite complex and full of information about enaction in other domains.

We know the sense of touch provides information about people, objects and the environment—the world around us. Less known, but equally important, touch gives us information about our internal environment—the world within us. I will describe how touch provides both objective and subjective information and show examples of how I have explored this capacity of touch through art-making.

Sight is swift, mobile and encompassing, but it is also inherently objectifying, distancing, disembodying, and has a point of view. Touch provides a more embodied, intimate, subjective, and diffuse mode of sensory exploration. The pervasiveness of tactile nerves throughout the body and the body's ability to move make touch a more flexible and complex sensory mode.

Touch takes place at several levels within the body: the skin, the muscles, and the joints. Each layer conveys different kinds of information about the world and about ourselves. Meeting at the level of skin gives us direct contact with things; we sense their reality in the concreteness, pressure and immediacy. On the subjective, internal side of the skin, the contact is felt as reciprocal: as we touch things, they touch us back, reminding us of our own reality and the conditions of that reality. They tell us we are firmer than sand, softer than steel, warmer than ice. This feedback from the environment is necessary for our sanity and well being for the duration of our lives. David Katz, perceptual psychologist, said in his book, *The World of Touch*:

> "We must give precedence to touch over all the other senses because its perceptions have the most compelling character of reality…Nothing convinces us as much of the world's existence, as well as the reality of our own body, as the (often painful) collisions that occur between the body and its environment."

Below the skin, there is knowing at the level of muscles and limbs through movement—kinesthesia or haptics. Static touch reveals very little. We know the shape, size and texture of what we touch by the range, shape and speed of our motions. At the same time our motions help us know and define ourselves by the form, direction and qualities of our movements, conveying where and how we are moving—fast or slow, with ease or tension. The quality of three-dimensionality produced by haptic perception induces feelings of solidity and substantiality, both in objects and in our bodies. This vividness extends to spatial relationships: the haptic experience renders space palpable and three-dimensional as we map it with bodily motions even though there ”nothing” there.

Deeper still, receptors in and around the joints of the body sense movement, weight and position. This inner sensing, proprioception, gives us access to our sense of movement as well as internal processes such as breathing, pulses, digestion, tension, pleasure, emotion, even memories.

Access to both subjective and objective realities is also integral to the experience of art. Artworks describe the worlds around us as well as the worlds within us. This capacity in both art and touch to reveal both subjective and objective dimensions makes touch deeply compatible with the experience of art. I began making tactile artwork for people with visual limitations, but soon discovered this capacity of touch to speak to both inner and outer conditions, and realized that touching art can be an enriching experience for everyone.

In our daily lives we are largely unconscious of the enormous amount of information we gather through the sense of touch. Because my sculptures are novel and complex, people are rewarded for paying conscious attention to what is usually unconscious. I make sculptures in which touch is integral to their structure, meaning and *aesthetics*—sculptures that are as coherent, compelling, and meaningful to touch as they are to see. I will show a few examples of my sculptures to convey my engagement with touch and what I have learned about the nature of haptics as a way of knowing. Through the years I’ve been working haptically, I’ve gathered comments from the people who touch my sculptures, usually blindfolded, which enables them to experience touch unmediated by sight. I discovered that the tactile version of an artwork often differs from the visual version. For example, people touching the same artwork reported remarkably different experiences. One woman, eyes closed, imagined she was in a forest, feeling trees and moss. When she opened her eyes she was astonished to see a geometric, metallic construction. A woman who had no sight said it felt hard and unyielding, making her think of factories and machinery and causing her to weep. Another woman, eyes closed, slid her hand along the outer surfaces of the walls. When her hand fell into one of the openings, the field of vision behind her closed eyes abruptly shifted from charcoal grey to deep black.

Seeing artworks only through a visual lens can distort and limit their meanings. The art historian Michael Brenson suggests in his article, “Memory of the Hand” in *Sculpture*, that Rodin, approached “in terms of his appeal to the hand, invites intimacy rather than the qualities of heroism and genius so visible to the eye and so often ascribed to his work.” Giacometti’s sculptures speak to the eye but also to the hand:

> “When you engage Giacometti’s figures with your hands, it is apparent that they have every bit as much to do with tenderness as they do with violence. To hold their shoulders, backs and heads is to ease the violence committed by the eye.“

In both examples, the tensions between tactile and visual impressions render the work more complex and layered. These contradictions widen the emotional territory revealed by the sculptures and reflect more fully the realities of being human, for we constantly live with such contradictions and tensions.

People who generate different versions of an object from sight and touch often realize that the sensory mode we use determines the nature of what we perceive. This is a profound insight. We assume seeing is believing, but when we realize how differently we perceive things through touch, we discover that sight is not so objective and that there are dimensions in plain view but still hidden. Because the linear nature of haptic perception reveals an object slowly, over time, some people grasp the extent to which they create the object through their actions and responses. The relatively objectifying simultaneity of sight creates the illusion that we know something, but there are many levels of knowing, and the superficial appearance is only the first level. Knowing something is a slow, gradual, deepening process, which hapticity encourages.

Touching allows people a direct, unmediated experience of an artwork, providing them with the authority of their own perceptual experience as a basis for knowing. The authenticity and concreteness of touch gives them confidence in their perceptions, cutting through physical and psychological distance and disarming the notion many people have that they don’t know enough about art to enjoy it.

Touch engages the body, drawing more of a person into the exchange than merely standing and looking. Contact, movement, and gesture generate more and different cues to memory, multiplying and enriching associations and meanings. The aesthetic experience is grounded in muscle and bone, gut and heart, rendering it closer to the original meaning of aesthetic, which is sensory. Because the body is involved, the stakes are higher. There is a feeling of risk, commitment and engagement that heightens concentration and deepens perception. Touching evokes vulnerability and danger as well as intimacy and affection. The sensations of touch can be compelling because of their closeness to, and resonance with, emotions, childhood, and sexuality.

The intimacy of touch radically alters people’s experience of an artwork; it becomes more fully known and owned. The reciprocity of touch creates the sense of being in dialogue with an artwork. This mutuality is the very essence of touch. Haptic knowing is horizontal, non-hierarchical and democratic. No longer simply subject and object, the art work becomes more like a subject—alive, sentient, responsive. Touching allows people inside the process by which the artist creates an artwork. People re-enact the artist's process and re-create the artwork for themselves. People feel connected to the artwork, to the artist, to the artist's process, to their own process, and to themselves.

These experiences arise at the unexplored place where someone's being touches the being of a sculpture. A door opens that swings both inward and outward. Such experiences challenge the hegemony of the self we locate in the eyes and brain. Because seeing and the seen carry so much authority in our culture, hapticity and enaction offer a needed balance to the dominance of the visual, evoking the bodily intelligence called for by Enaction. This work offers a small antidote to the increasingly disembodied lives our technology and our education are shaping, opening us to ways of knowing and being that we must reclaim.

Biography

A sculptor, painter, papermaker, writer and educator, Rosalyn Driscoll has for the last fifteen years made sculptures which may be touched as well as seen. Her deep exploration of aesthetic touch takes the form of making and exhibiting tactile sculptures, gathering viewer reactions, following research in tactile/haptic perception, working with scientists, engineers, artists and people with disabilities, lecturing and teaching workshops, and writing a book, *Whole Body Seeing: Touch in the Visual Arts*.

Figure 1. *Pandora's Box 8*, Rosalyn Driscoll, 2006


Figure 2. *Pandora's box 6*, Rosalyn Driscoll, 2004

RefleCT/Xion
from enaction to "daily enaction"

Ricardo Atienza Gabriel Bérubé Aurore Bonnet Damien Masson juL McOisans
Laboratoire Cresson
UMR 1563 CNRS
MCC «ambiances architecturales & urbaines»
Ecole nationale supérieure d'architecture de Grenoble, France
Email: {ricardo.atienza;aurore.bonnet;damien.masson;julien.mcoisans}@grenoble.archi.fr

1. Introduction

The concept of enaction deals with a bodily experience, an action knowledge. The body-in-movement is not a pure figure of action. It is, first of all, interaction with our environment that reveals us the limits of our own bodies and also the qualities of the space that surrounds us.

The traditional approach to enaction was by notion of "interface", however the purpose of this approach is not to point out the tension that exists between body and space but rather the possible devices and mechanisms that can act upon this relationship. This interface comes in the form of any tool manipulated by the hand, to acquire more complex technical expressions, thanks largely to computer equipment. The concept of enaction is thus translated in terms of inter-relation between man and machine. This interface (man / machine) is dedicated sometimes to emphasize certain qualities of a space but also to the creation of a virtual space that is superimposed on the physical context.

In contrast to these 'indirect' modes, there exists a primitive and direct enaction, that of the bodily experience of space. The interrelationship of man and machine is replaced by the simple interaction between man and space. By moving, touching, listening, watching and feeling, the body 'understands' and decodes knowingly and reads extensively its surroundings. Browsing an area is the first condition of any appropriation. Some senses allow us 'to touch by distance' but others require a direct contact - our skin, the last frontier before the world, is an essential gatherer of information about our environment. The body's motor acts offer us a dynamic reading of our environment where its sensitive variations (thermal, aerolics, light, sound, smell…) stimulate our perceptions and reveal the various possible readings of a space.

The concept of walking is a fundamental idea of this proposal: walking to and within a space; walking senses in motion, walking as a daily experience that teaches us balance, orientation, intuition, intent and our interpretation of space. Our 'ordinary' experience is composed of an infinity of enactive fragments. Our daily life is a continuous apprenticeship by the body of its surrounding space.

The various works of the exhibition propose to reveal the ordinary experience of walking to the visitor, to make one aware this ordinary experience of interaction between body and space.

2. The experience of everyday life

Despite being a widely-used concept, enaction is an important characteristic of everyday life. Thus our approach is aimed to focus on this very ordinary experience of space that one might call "daily enaction".

Our proposal uses some ideas inherent in a broad range of CRESSON's research dealing with "ambiance". The field of architectural and urban atmospheres, which crosses disciplinary boundaries by surrounding simultaneously perceptible social and built space, focuses on the ordinary perception of the architectural and urban environment in order to understand different manners of "dwelling" in it. Therefore, our theoretical approach, which underscores the installation created at the Fort de la Bastille, focuses on the sensory interactions between human beings and the "world" they experience daily.

These interactions are numerous and encompass the perception of oneself, of others and of the built and non-built environment simultaneously. Questioning such multiple interactions as well as everyday sensory experience provides a way to understand enaction in its ecological perspective by focusing on the "middle ground" that exists between built forms and human behaviour. One way of revealing this 'in-between' is proposed by James J. Gibson in his pragmatic notion of "affordance" which deals with the possibilities for action allowed by space and objects, however we have chosen to focus on the "effects" that qualify it.

The "intermediary concept" [2] of effect is a useful means for describing what happens between phenomena and perception. By emphasizing the fact that people co-produce effects while they are acting in an environment and by considering the sensations (and not only the perception), this notion rejects any consideration of the "middle ground" phenomena and perception in terms of causal relations. Thus we have chosen to focus on the sensory effects that structure an individual's perception of space in order to question their "daily enaction".

Erwin Straus [7] proposed to understand perception by defining it by way of the dialectical link that exists between sensations and movement. We have chosen to focus our attention on the act of walking as a way to reveal enaction in the mundane experience of space. Subsequently our purpose is to turn the ordinary into the extraordinary through the common act of walking. For this we propose to look at the "ambiance" of the space, then our process will focus on the possibilities for subtle alterations, which would help a visitor to become aware of their acts. We then propose to implement some very simple systems, which will both alter some sensory properties of the space and react to the visitor's presence. "Reflection", understood both as visual reflection and mental reflection, is the principle that guides this implementation.

3. Suspended space, textured space, reflecting space…

Practically, working on "daily enaction" through the principle of reflection supposes to imply some means that reveals and "reflects" the interrelation between the body and the space. This interrelation reveals itself through sensible effects; with the aim of becoming aware of those ordinary effects we need to emphasize them, to accentuate what is simply usual in our daily experience.

In order to call attention to those ordinary effects we manipulate the physical context supposed to produce them: perception of space and time will be slightly deformed while the body moves in it. This kind of "amplified" or even "interfered" consciousness is induced through three different actions: working ground-textures, modelling the space through suspended textiles and light materials, reflecting and manipulating image and sound of corporal movements.

The experience of the ground textures mainly reveals the physical aspects of walking: learning through walking. By covering the ground with different materials (pebbles, dead leaves, strips of wood or metal, etc.), a consciousness is evoked through the kinesic experience itself, but also through the 'sounding' of the visitor's steps. Another aspect of the exhibition, and the visitor's movement within it, is suggested by the textures (chalk) they leave in their path by way of footprints. Footprints, rubbing, creaks, rumblings and echoes generated by steps invite the visitors to investigate the relationship with the ground on which they walk. The textural effects experiences tend to reveal it as a fundamental substance of our environment.

The implementation of suspended textiles is to work on the organisation of the space in different ways. Partitioning without creating solid, rigid walls, adding depth and various levels in which the visitors immerse themselves. The density, transparency or reflectivity (tulle, lycra, rhodoïdes, mylar, etc.) allows successive dimensions of projection and lightning to alter the perception of the space itself.

Textiles are also used as supports to projection of real time captured images of the visitors. These projections are based on the phenomenon of reflection; reflection of ourselves, of our body-experience, reflection of the image of our movements. But in order to emphasize this self-experience, this reflection is worked as a deforming or moving mirror : Larsen, delay, distortion, echo, … describe some of the effects used. The idea of these effects (visual, audible, tactile, olfactory…) works according to the principles of incongruity and amplification of reality, allowing the visitor to understand that their actions have an effect on what they perceive, beyond the immediacy of the principle action / reaction.

4. Walking enaction for a "mise en abyme"

Contrary to this recent perspective about the definition of enaction, which puts forth the idea of an interaction between man and machine, we opted for a more literal approach that combines people and their environment. The setting, a place we can experience in our daily actions, is projected and amplified by simple means. Conceived to be a link through different enactive installations, our aim was to try to construct a kind of "enactive path" between them.

Working in the CRESSON laboratory principles, that advocate an interdisciplinary, qualitative approach, *in situ*, the notion of atmosphere ("ambiances") becomes a base in the analysis and design of the project. The idea is to reveal the slight changes incurred by some of the trivial actions of the visitor in order to raise awareness of these phenomena. Passing through the space, senses, searching for information, will be exacerbated by the surprising and mysterious data that they acquire. Then, walking becomes a new means for the discovery and the understanding of our environment.

With only a few equipment and some little technological devices that permit us to work with sonic and visual effects, we have found a simple way to introduce the visitor into the "world of daily enaction". In their strolling, the visitors will find a place that make them feel like they are, at the same time, actors and a spectators of their surrounding by their simple ordinary movements.

References

[1] Amphoux P. , Thibaud J.-P., Chelkoff G. (eds.), *Ambiances en débats*, Bernin: À la croisée, 2004.

[2] Amphoux P., L'observation récurrente : une approche reconstructive de l'environnement construit, *in* Moser G., Weiss K. (eds.), *Espaces de vie. Aspects de la relation homme-environnement*, Paris: Armand Colin, 2003, p. 227-245.

[3] Augoyard J.F., *Pas à pas*. *Essai sur les cheminements quotidiens en milieu urbain*. Paris, Le Seuil, 1979.

[4] Berthoz A. *Le sens du mouvement*, Paris: Odile Jacob, 1997.

[5] Gibson J.J., *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin, 1979.

[6] Gibson J.J., Le système haptique et ses composantes, *Nouvelles de Danse*, 2001, n°48-49, p. 94-120.

[7] Straus E., *Du sens des sens*, Grenoble: J. Millon, 2000 (1st ed. 1935).

L'énaction comme éclairage d'une pédagogie des arts dynamiques

Daniel Barthélemy
EESI Angoulême, France
Email: d.barthelemy@eesati.fr

Résumé

L'Ecole Européenne Supérieure de l'Image est une école d'art engagée depuis sa création dans un questionnement relatif aux arts numériques. Dans ce cadre, un accord passé en 2002 avec l'ACROE a permis aux étudiants de se confronter à l'univers des recherches menées par notre partenaire.

L'émergence du concept d'énaction nous apparaît aujourd'hui comme un vecteur particulièrement approprié à rendre perceptibles les enjeux de redéfinition de notions de base telles que celles de supports, de matière ou d'interaction qui occupent l'ensemble du champ des arts plastiques.

L'expérience menée au cours de ces échanges a permis à l'EESI de développer avec son partenaire une compréhension plus profonde des enjeux de recherche et des implications esthétiques que posent les nouvelles technologies tant dans leur développement que dans l'enseignement auquel les écoles d'art auront à faire face dans les années à venir.

Depuis près de six ans l'EESI a engagé un projet d'enseignement autour de l'environnement Cordis-Anima. Cet engagement s'est concrétisé par de nombreuses interventions des chercheurs d'ACROE, la mise à disposition de l'EESI des logiciels MIMESIS et GENESIS, l'organisation d'une quinzaine d'échanges sous forme d'ateliers à Angoulême et Poitiers ainsi que l'accueil d'étudiants à Grenoble. En 2005, le projet ARCAD, soutenu par le Ministère de la Culture et de la Communication a permis l'intensification et la pérennisation de ces échanges.

Cette expérience nous a permis d'approfondir un peu plus la façon dont peut être dispensé et perçu un tel enseignement auprès de la population des étudiants actuellement accueillis dans les écoles d'art.

Le choix de cette expérience était dicté par le caractère particulier de l'environnement CORDIS ANIMA dont l'un des principes fondamentaux est le principe multisensoriel des objets traités : visuels, sonores, haptiques. Pour un enseignement raisonné des problématiques du multimédia il constituait un objet d'expérimentation d'autant plus significatif que son formalisme propre imposait une approche singulière du mouvement qu'il aborde non plus de manière phénoménologique, mais causale. Par ce principe, il contribue, comme d'autres environnements privilégiés dans notre réflexion à installer une position différente, critique, des environnements conventionnels.

Ainsi, la contribution dans ce projet des enseignants du multimédia, des espaces sonores et d'un ingénieur de recherche a permis le déploiement d'interventions en mesure de présenter les diverses formes des objets traités par cet environnement.

A l'égard des enjeux que pose l'évolution nécessaire de notre enseignement confronté aux nouvelles technologies, le concept d'Enaction est arrivé comme un éclairage salutaire, permettant de fédérer des notions disparates, orphelines, toujours réductrices, que nous sentions jusque là impuissantes à tenir ensemble une intuition de présence dont on voulait entourer les objets virtuels dans la présentation que nous devons en faire à nos étudiants. L'ensemble Cordis-Anima a pu constituer pour nous et pour une trentaine de nos étudiants, l'intercesseur radical, le maillon manquant dans la chaîne instrumentale d'où émergent les instruments numériques.

Cela dit, ce type d'approche des objets numériques se confronte à un environnement en profonde mutation, qu'il s'agisse de l'évolution structurelle des écoles d'art tout autant que de l'enseignement des technologies numériques, elles-mêmes en continuelle gestation.

La forte médiatisation, dans le milieu, des expériences développées ici et là conduit les étudiants à disposer rapidement d'une connaissance de plus en plus accrue de la palette d'innovations auxquelles ils peuvent consacrer leurs études. Cependant, la nature et les enjeux de ces diverses initiatives reste encore d'une lisibilité insuffisante. Il semblerait qu'un enseignement complet nécessiterait actuellement pour un étudiant désireux de suivre les recherches menées dans les différentes écoles, de suivre plusieurs laboratoires au cours de son cycle d'études.

L'ordinateur, tel que l'a imaginé l'industrie informatique afin d'apprivoiser rapidement un nombre d'usagers croissant a donné naissance à des instruments totalement inédits pourtant présentés comme de banales évolutions des instruments familiers du grand public. Derrière l'écran de l'ordinateur domestique, derrière les icônes et les divers raccourcis dont il semble tissé, toute une culture algorithmique, toute une richesse de concepts se sont développés qui ont profondément modifié non seulement l'ensemble des médias, mais aussi tout l'univers de nos représentations. Plus de vingt ans après ce mouvement, il apparaît indispensable de permettre aux jeunes artistes de s'approprier cette culture et de pouvoir affirmer dans ce nouvel horizon, leur attachement historique à maîtriser leurs propres outils.

A présent, la rencontre de nos étudiants avec les objets physiques, tels que de tels processus les mettent en oeuvre pourrait se résumer ironiquement dans cette image que rapporte Annie LUCIANI évoquant une étudiante qui souhaitait modéliser « une lumière à retour d'efforts ».

Les formes de pensée que propose l'environnement tel que celui développé par ACROE se heurtent encore à un ensemble de difficultés de mise en œuvre dans une Ecole d'Art ; environnement matériel, maintenance et accessibilité, apports et développements théoriques appropriés. Mais surtout, une telle initiative souffre encore de la nécessité d'une reconnaissance accrue pour l'ensemble des écoles d'art, du caractère fondamental de tels questionnements. Tels qu'ils sont formulés, les fondamentaux de nos enseignements ne permettent pas à nos étudiants de saisir la nature de ces enjeux.

L'accueil d'une nouvelle expérimentation semble activer un ratio destiné à devoir choisir entre un environnement intuitif et convivial indépendamment de la simplicité des formes qu'il peut générer ou un autre, plus aride dans son approche conceptuelle et technique, mais mettant en jeu des questions d'une portée plus fondamentales.

Si ce ratio peut relever d'une forme de « Beaux-Arts attitude », la culture de l'intuition à laquelle l'industrie informatique a habitué ses jeunes utilisateurs n'y est pas non plus totalement étrangère. Ainsi le passage par de nouveaux modes de pensée n'est pas toujours perçu par nos étudiants comme le chemin nécessaire d'une plus ample compréhension des processus, mais relégué à une forme d'archaïsme, de désuétude qu'ils préfèrent abandonner à leurs aînés.

De plus, dans le monde plasticien un tel environnement est frappé d'une suspicion encore persistante. D'un côté, il lui est reproché un manque d'interactivité, de présence physique et d'un autre un trop d'instrumentalité, de dureté mécanique. Ce préjugé pourrait conduire à se demander s'il ne s'agit pas, au bout du compte, du seul et même côté d'un manque plus profond ? De ce manque précis d'où émerge le besoin de création qui constitue le propre de l'art ?

Des préjugés

Il semble que notre expérience a confirmé la présence de formes diverses de résistance dont il reste difficile d'identifier l'origine et les enjeux qui semblent correspondre à plusieurs systèmes de pensées auxquels notre projet s'est retrouvé confronté.

L'accueil d'une pensée d'origine scientifique, aussi proche de questions culturelles ou esthétiques soit-elle est encore perçu comme une tentative de surplomb des problématiques développées dans le champ de l'art. Cet état de fait dans l'enseignement supérieur ne sera probablement pas réglé tant que les enseignements secondaires resteront scindés comme ils le sont encore par l'étanchéité des filières qui préparent nos futurs étudiants à l'enseignement supérieur.

De son côté, le monde de l'art contemporain valorise actuellement le pouvoir de mettre en lumière un concept avec une minimum de technicité. Les enseignements liés à l'infographie dans les filières artistiques ont constamment rencontré cette opposition depuis leur apparition voici près d'une trentaine d'années.

Il faut rappeler que le concept même de machines ou d'interfaces “intuitives” - *Think different* - a largement contribué à écarter toute la culture technicienne ou procédurale, pourtant inhérentes à toute l'histoire de l'art.

Le regard que nous ouvrent les approches actuelles de l'univers scientifique bien qu'il bouleverse chaque jour un peu plus notre conception de la matière, du mouvement ou des relations entre les corps, nécessite d'apprivoiser un ensemble de concepts nouveaux, de systèmes de pensées qui semblent encore trop étrangers aux pratiques courantes, et se trouve de ce fait souvent rejeté. L'évolution de l'art n'invalide généralement pas les concepts précédents comme peut le faire la conscience scientifique. Ainsi des modèles de représentation de la matière ou du mouvement tels qu'en ont produit le moulage ou le cinéma restent des patterns dominants qui parviennent parfois à obscurcir totalement l'hypothèse de nouvelles formes de représentation.

Le climat incertain des débouchés aux études ne permet probablement pas à nos étudiants de trouver la sérénité nécessaire à faire les choix essentiels dans leur cursus d'enseignement. Tout le parcours d'un étudiant confronté aux nouvelles technologies est jalonné d'acquisition de savoirs faire dont il ne peut pas être en mesure d'évaluer la pérennisation ou l'obsolescence prévisibles. Ainsi il sera contraint de faire des choix plus souvent dictés par les opportunités immédiates que par une véritable analyse des enjeux à moyen ou long terme, ou du caractère fondamental des acquisitions qu'ils apportent.

Cela vaut également pour les différents apports théoriques tels qu'ils sont actuellement ménagés dans les programmes de leurs études et tels qu'ils pourront prendre place dans l'ensemble des apports théoriques.

Images de la matière

L'irruption d'environnements issus du monde de la physique dans le champ des disciplines plastiques s'est trouvé confronté au paradoxe d'être en même temps trop 'technique' et 'trop virtuel'.

Ce paradoxe ne touche pas seulement l'environnement matériel, relativement aride que constituent des stations UNIX auquel les étudiants sont rarement confrontés. Il atteint la notion directe de matière tout autant comme discipline qu'en tant que principe même de la plasticité.

Si le physicien envisage la matière comme un espace dans lequel s'agencent des particules, ce regard questionne pour nous non seulement la notion rassurante de matière en tant que solide, qu'élément plastique, mais également tout ce que les plasticiens des nouveaux médias viennent d'élaborer comme concepts relatifs à l'interactivité.

Le regard qu'ouvrent à l'imaginaire les représentations que nous apportent l'échelle nanométrique rencontre de grandes difficultés à franchir les enceintes de nos certitudes. Pourtant il nous apparaît immédiatement à l'esprit que l'expérimentation physique de telles échelles de représentation rencontrera dans notre imaginaire une résonance spontanée en mesure de générer de nouvelles formes d'interaction avec le monde.

Images de la recherche et vice versa

Les impératifs de développer des activités de recherche sont à présent un double enjeu très vivace dans les Ecoles d'Art. Pour une école telle que l'EESI engagée vers les nouvelles technologies depuis sa fondation, ces enjeux font partie intégrante de ses objectifs structurels.

Les nouvelles technologies sont par définition un domaine où les savoirs, les pratiques et les enseignements sont en perpétuel renouvellement. Le second enjeu plus décisif est celui d'une reconnaissance d'un enseignement « supérieur » relatif aux accords de Bologne dans le cadre des processus d'harmonisation des enseignements dans la communauté européenne.

Il va de soi que cet enjeu se confronte aujourd'hui sur les « images » que nos Ecoles d' Art entretiennent à propos de la recherche, de la science, de la matière ou de la mécanique.

Franchir le stade de ces images pour accéder à un peu plus de réalité, nécessite une véritable politique de circulation du sens à l'intérieur de nos départements, de nos équipes, de l'ensemble des établissements d'enseignement artistique engagés vers ce défi.

Pour notre part, nous poursuivons et développons ces expériences car nous constatons au fil des années que nos étudiants évoluent, nos politiques d'enseignement également, mais surtout qu'en fin de compte, les principales difficultés d'intégration évoluent clairement dans le sens de leur résolution.

Projections

L'analyse de la nature des questions rencontrées durant notre expérience nous a conduit à définir de nouveaux modes d'action plus appropriés qui feront l'objet de nos expériences à venir.

Il semble urgent que le projet soit rapidement relayé par de jeunes acteurs. La génération qui le porte appartient à une première approche de l'informatique et il est paradoxal que les concepts défendus au travers de l'historique de cette recherche (pourtant en prise directe sur des domaines technologiques de pointe), soient de ce fait identifiés par de jeunes étudiants artistes comme des problématiques anciennes, soi-disant révolues.

Il apparaît dans la plupart des établissement du second cycle de l'enseignement secondaire que les filières en place conduisent les élèves à choisir entre arts plastiques et sciences. Les Arts plastiques comme « options lourdes » ne sont possibles que dans les filières littéraires de cet enseignement. Il est bon de noter que cette alternative n'existe pas dans de nombreux cas à l'égard de la musique et qu'ainsi, une distinction fondamentale d'approche des disciplines techniques ou scientifiques sera alors développée selon que l'on aura choisi l'un ou l'autre de ces arts.

Au moment où les nouvelles technologies, telles que nous les avons décrit plus haut, permettent d'amener la création visuelle vers une nouvelle approche sollicitant la totalité multisensorielle de nos perceptions, il semble urgent de devoir reconsidérer ces modes de cloisonnement de l'enseignement.

L'apparition des premiers outils a imposé, par les moyens rudimentaires dont disposait l'homme de focaliser les concepts d'expression selon les différents canaux perceptifs, tels qu'ils restèrent longtemps catégorisés par la physiologie. Il a fallu attendre de plus lointaines découvertes pour estimer plus précisément comment chacun de ces canaux sensoriels concourent ensemble à notre compréhension globale du monde. Il faudra encore attendre de nombreuses années avant de reconnaître que la somme de ces canaux ne constituait pas un tout, qu'il doit rester du perceptif non localisé, quelque chose comme la musique des formes ou l'image des sonorités, toutes formes de goûts indécelables individuellement, mais indiscutablement présents dans notre environnement psychosensoriel et jouant entre ces canaux par d'insaisissables relations.

L'environnement Cordis-Anima, on l'a vu , se place au cœur de cette recherche d'une totalisation sensori-motrice et il nous permet de nous questionner au-delà de la façon cosmétique dont on plaque fréquemment le son sur l'image, l'interaction sur le contact, l'illustration sous le texte.

Mais à ce moment, la question se pose qui mériterait peut-être plus ample attention. Comment émergeront de futurs artistes en mesure d'exprimer cette approche globale, enactive auprès d'individus que de semblables cloisonnements auront également formatés ?

Perspectives

Parallèlement à ce projet plusieurs événements plus ou moins connexes ont survenu qui en ont renforcé l'évolution d'un esprit de recherche dans l'EESI. Notamment le développement du projet « *Sliders* » qui a mobilisé une bonne partie de l'équipe pédagogique et renforcé l'image et l'affirmation de l'EESI dans les nouveaux médias.

D'une manière générale il est évident que les impératifs d'évolution vers le cycle LMD ont stigmatisé cet esprit de recherche encouragé par le soutien de notre Direction.

Mais aussi, cette expérience nous a incité à établir davantage de contacts auprès des filières de l'enseignement secondaire afin de mieux sensibiliser nos futurs étudiants aux problématiques qui les attendent à présent dans l'enseignement dispensé dans les écoles d'art.

Les échanges initiés par le projet entre Gilles BOLLAERT, ingénieur de recherche dans l'EESI et développeur de *Synthetic Video* avec l'ACROE a permis de mettre en œuvre une dynamique de nouveaux développements dans ce logiciel actuellement développé à l'intérieur de l'école.

Ces développements ont fait l'objet d'une série d'ateliers et de travaux durant toute l'année universitaire 2006-2007 et vont prendre la forme d'un Atelier de Recherche et de Création (Atelier de Recherche et de Création *PIXMEM*) sur l'ensemble des deux sites de l'EESI .

D'une manière générale, le rapprochement de l'EESI et de l'ACROE durant le déroulement de ce projet a renforcé notre conscience du dynamisme de la pensée scientifique.

La question est actuellement de développer chez nous la mesure dans laquelle les recherches scientifiques contribuent à questionner sans les nier, les invariants conceptuels de notre pratique artistique.

Un profond travail théorique doit à présent se poursuivre, dont les apports du concept de l'*Enaction* éclaireront différemment notre pratique. Les enrichissements conceptuels qu'il apporte par son aspect pluridisciplinaire devraient lui permettre de jouer un rôle décisif pour l'avenir de l'enseignement des applications plastiques des arts dynamiques.

Musical Creation Process and Digital Technology
"the Supra-Instrumental Gesture"

Claude Cadoz
ACROE and ICA Laboratory, France
E-mail: Claude.Cadoz@imag.fr

Abstract

Considering artistic creation as an emblematic enactive process and the artistic creation process as deeply linked with technology, we propose a conceptual framework allowing to give signification to the concept of artistic creation tool in the context of digital technology. We start from a very simple (theoretical) situation, which is a kind of primordial musical instrumental experience where we characterize the main phases of the creation process under an enactive point of view. Then, we discuss briefly some needs to which technology evolution brought certain solutions and the new functionalities it introduced. After that, entering in the "new technology" era we show in what the introduction of the digital technology in the field of artistic creation is not simply an evolution but, though not yet accomplished, a deep revolution. Finally, through a brief presentation of our own research and its correlated musical creation activity in the laboratory, we introduce two important concepts: the first concerns the use of the mass-interaction physical modeling paradigm for the musical macrostructure creation, the second, closely related, is that of "Supra-Instrumental Gesture". Both correspond, in the case of musical creation, to an extension of the enaction point of view from the multisensory-motricity level to the scale of musical composition.

1. Introduction

Evoking artistic activity, we can start by posing the problem in terms of communication. Indeed, there is at least, at a given moment and a given place, between several people, some objective intermediary produced or influenced by the ones, and perceived by the others. And at the minimum, this intermediary is a physical phenomenon, for example an acoustic wave, or a light flow.

Of course, a lot of other conditions are necessary for the existence of artistic creation situations, but focusing on the natural conditions in which the human being can produce or modulate phenomena for the senses and perceive the phenomena given to its senses, before introducing any technology, we can bring in the enactive considerations.

Let's recall it through four primary remarks:

1) There is a dissymmetry between the ways to produce and the ways to perceive: Hearing is able to treat a much larger variety of acoustical phenomena than what the voice can produce. The human being is the source of the acoustical energy the voice produces and which reaches the hearing, while an external source of energy (light) is necessary in the case of the sight. Body behavior, face expressions and purely semiotic gestures [1] are visually perceptible, but ergotic [1] gestures phenomena (mechanical forces, deformations, displacements…) are not propagated and suppose a direct physical contact to be perceived by the tactile / haptic sense.
2) There is no perception without production and no production without perception. Production / perception processes are loops. For example, in a vocal (musical or not) communication, the following flows are concerned:
    a. Vocal production to hearing of emitting subject.
    b. Vocal production to proprioception of vocal organ of emitting subject.
    c. Vocal production to hearing of receiving subject.
    d. Body and face action-movement to proprioception of emitting subject.
    e. Body and face action-movement to sight of receiving subject.

    Of course, there are loops at higher levels, in natural communication situations, where the subjects are more than two, and where the phenomena are for example complex structured audio stream related to speech or song.

3) Not more than it is possible to completely separate production and perception, it is possible to thoroughly separate our senses, or our way of production. This is the reason why enactivity is strongly associated with the idea of system of “multisensory-motor loops”.
4) And finally, it is necessary 1) to distinguish between “production” (in the sense of producing or modulating phenomena for the senses) and “action” (in the sense of physical action on, or by, a material object), and 2) to consider that, again, they cannot be absolutely separated.

While technology intervenes, changes concerning dissymmetry can occur: for example, in the case of music, a musical instrument can be viewed as a means to increase the variety of acoustical phenomena that the human being can produce, as well as a way to allow the ergotic gesture to propagate. In the same time, a musical instrument can extend the spatial era where the communicational process can take place, since, for example, we can ear its sounds at a longer distance than those of the voice. However, these advantages are not free. When technology intervenes, there is a transaction, a kind of “deal” and the new situations are at the same time more than, less than, different from, the previous. For example, according to its material constitution, the musical instrument has its specific affordance, imposes specific gestures, and has its own vibration modes. In fact, it reveals the action of the instrumentalist by revealing itself, indissolubly, adding, losing, and transforming certain things.

This is this question of “transaction with technology” that we discuss in an overflight along these few pages, from the “primordial” instrumental situation to the advent of the digital technology in the field of the musical creation. The question being actually to try to characterize the strong discontinuity between the “classic” and the “contemporary” (digital) technology. Recalling the main topics and previous results of the research in our laboratory, we will then propose some prospects which, while being established on a major form of continuity, brings new latitudes truly possible only with the computer and the enactive systems.

2. Musical creation process and “Classical technology”

2.1. The “primordial” instrumental situation

The musical instrument is the first stage of technology in music. We define it, here, as a material object with which we can physically interact through ergotic gestures and in which, as a consequence of this interaction, take place vibrating processes that produce acoustic phenomena. So, a musical instrument converts a gestural energy into acoustic energy, a physical **action** into an acoustical **production**.

We said above that the musical instrument enlarges the variety of sounds that the human being can give to hearing, at the same time as it allows the (ergotic) gesture to propagate or, in other words, the gesture to reach the ear. Taking into account the laws of the physical universe, it can do it only according to certain types of protocols and thanks to certain categories of physical process. This has as consequences 1) that the physical structure of a musical instrument presents a prototypal chain of physical components, and 2) that the gestures we can do respond to a precise typology (*excitation / modulation / selection* gestures [2]), and has specific morphology (hitting, rubbing, scraping, bowing, blowing, plucking, etc [3]).

What is very important in this “primordial” instrumental situation is that there is a continuum of energy from the gesture to the sound. So, quantitative (amplitude), but also qualitative (timbral aspects, etc.) properties of the heard sounds are directly linked to the human sensori-motor capabilities, the physical laws of the universe through the material forms that the instrument can adopt, and through the physical contingencies of the man-instrument system.

The creation and communication processes (envisaged not only at the level of the individual, but in their historical and social dimensions) are, then, specifically determined and we can, again, adopt the enactive point of view to characterize them.

Let’s start with an hypothetic primary situation where we are using for the first time an object about which we just know that it is a musical instrument, but nothing about what gestures we can apply, what sounds it can produce and what is the link between the gestures and the sounds. The only possible attitude in this situation is to explore, i.e. to act in order to perceive, and to perceive in order to act. Starting from scratch (which is never from scratch), the exploration is done according to strategies determined by our physical (biological) capabilities, the mechanical properties of the instrument, the physical possibilities and properties of the man-instrument system, the intrinsic affordance of the instrument, and our previous… enactive knowledge, i.e. previous situations where some aspects of the present one were similar.

Doing this exploration, something new happens, i.e. production, for example, of sound events, which we don’t know before. So we get a new enactive knowledge linking the actions and the produced effects. But here we have an actual creation process, in its simplest form but in its very essence. Creation because we brought to the existence something which did not exist (for us) before. Is this corresponds to something we imagined before without being able to give it tangible (perceptible) form? May be yes, may be no. The fact is that now we know it and how to produce it. Is this corresponding to something, which is relevant? Beautiful? Meaningful? According to what judgment? If we judge that this is the case, then, the most important is that we can now propose it to the perception of other people, as often as we (they) wish it. The experience of other people with this event may correspond to something completely banal or completely new. In the last case, we can say that the individual creation is also available at a more collective level. Let's emphasis here how enaction and (artistic in this case) creation are deeply linked. Quite the same concept, but at different levels. Acting and perceiving, we are always creating, and creating we are always perceiving and acting.

Having sufficient (direct or indirect) instrumental experience, we may become able to achieve such exploration and creation process to a certain extent at a purely mental level, without any actual sensori-motor concretization. It is interesting here, without enter in details, to envisage a loop representation of a more global creation process where, at an upper level, the poles of the loop are the actual sensor-motor loop and the "mental loop" themselves.

2.2. Musical notation and "classical" composition

Painting can be presented like music, with notions of material intermediary, gestures, production and perception of phenomena. However the things are quite different. For example, the ergotic gesture is not transformed in energy for the senses (sight). But a major difference is that the material intermediary (playing the role of the instrument in music), i.e. brushes, canvas, etc. is at the same time a physical object that allows to preserve the result of the creating process over time and space. Painting, which is an "autographic" art [4] has *de facto* a natural spatio-temporal extension, beyond the moment and place where the creation process occurs.

In music, the sound phenomenon is fugitive. It disappears as soon as the action on the instrument ceases. This is the reason why music is "allographic" [4], i.e. it needs, beside the instrument and the human memory, an external physical permanent support in order to get a wider spatio-temporal (socio-historical) extension.

The music notation was invented during Antiquity, by the Chinese and by the Greeks, but it actually developed in Occident only from the Middle Ages, from the *neumes* [5] which were graphical signs allowing to transmit songs associated to religious texts from masters to disciples. The development of occidental music is strongly determined by the musical notation which main principles are very stable until contemporary period.

But musical notation, which is not dissociable from technology (of its supports, markers, etc.), supposes a certain relation between "musical facts" and graphical symbols on the support. The questions are then 1) what are the "facts" that are submitted to notation? And 2) what are the properties (affordances) of the graphical symbols themselves, in relation with what they represent, and intrinsically?

In fact, the relation between the musician and the score, through musical notation, must be considered itself as a specific creation process. It is of a completely different nature than the instrumental process. And more: the creation process – "composition" - in occidental music, is of a specific nature while combining, in particular contingencies, instrumental experience and formal representations.

3. Contemporary technology

3.1. The early new technology situation

Traditional musical notation is unable to transmit anything concerning the fine properties and details of the sound as our hearing is, however able to discern it. Indeed, in traditional notation, the sounds are defined only as pitch, duration, intensity and timbre (the last being evoked by reference to the instruments).

One of the important technological revolutions in music, at the XXth century, is due to the advent of sound recording (T. Edison and C. Cros invented the phonograph in 1875, the magnetic tape recording was invented in 1931). It gave rise to the "Musique Concrète" in 1948 [6].

It is simply obvious that the musical creation process, in "Musique Concrète" or in "Music for Tape", is completely different, while sound events, once fixed in a permanent support (the magnetic tape), became "sound objects" (*Objets Sonores*) according to the famous term from Pierre Schaeffer. The tape-recording, in fact, beyond the possibility to "capture" the sound and to "replay" it identical to itself indefinitely, establishes a specific correspondence between time and space: a given duration of the sound corresponds to a proportional length of the tape. Then, transposing our acts from instrument playing or score writing to tape cutting, inverting, pasting, etc. we get a way to work, "by proxy", on the temporal dimension, by working on the spatial one. The creation process is of course very new and includes the possibility to transform the sound, to go through time, and even in its reverse direction.

But in the same time (here are the terms of the transaction), the music became "autographic", that is the height! Indeed, the tape is the permanent object that gives to the music (for tape) its spatio-temporal extension, and (but) it is no more necessary (possible) to note it. And as a consequence to "compose" it, … in the traditional sense.

3.2. The coming on of computer and digital technology

Less than ten years after “Musique Concrète” and “Electronic Music”, computer entered in music domain with Automatic Composition and also Digital Sound Synthesis [7].

Automatic Composition, which used computer to mechanically perform formal processes corresponding to rules that “classical” musical composition used, became soon “Computer Assisted Composition” (CAC). This also inaugurated a very new type of creation process by introducing a kind of meta-level of dialogue: between the composer and his system of rules.

Digital Sound Synthesis, as for it, is closer to the senses. It could be introduced thanks to the invention (at the Bell Laboratories, in USA, in 1956) of the Digital to Analog Converter. This is by this device that a link between the electronic representation and manipulation of numbers in a computer and the ear (through electronic amplifier and loudspeakers) could be established. However, the important way that Max Mathews and Jean-Claude Risset [7] opened, followed by a lot of people, did not give a place to the instrumental (ergotic) gesture. This was of course because of the duration of the calculations, which did not allow producing each second of digital sound signal in less than one second. But a deeper reason was (and remains, despite the advent of real-time), that the algorithmic processes to calculate this enormous amount of samples were conceived as formal rules derived from the signal theory and processing. While gesture, when performed, has, cognitively only very little to deal with the Fourier Serial Decomposition or Transform.

The creation process within the digital sound synthesis is founded on the bridge that Pierre Schaeffer anticipated [6], and that the early works of J.-C. Risset (on the digital synthesis of trumpet sounds, for example) allowed him to theoretically well pose, between the physical description (whatever it is) of a physical phenomena, and what our perception does with it. The Psychoacoustic became then a constitutive loop of the musical (digital) creation process.

It is fundamentally new and powerful. It allows, as J.-C. Risset said, to compose not only the sounds together, but also the sound itself, in its intimate structure. It introduced also a revolution in the technology for music, allowing giving an objective and completely formalized counterpart of the instrument through models described in the MUSIC V language, and gathered in shareable catalogues [7].

We would like now to take support on this overflight of what we consider as important transitions in the technology for the music, to justify the orientation of our research, since its starting point in our laboratory.

4. Computer as a system for representation

4.1. The multisensory interactive simulation of physical objects

We started our research considering 1) that the instrumental experience, as we defined it above, is primordial in the musical creation process, not only envisaged at the individual level, but at the scale of the human evolution, and 2) that the digital technology revolution is likely to make us reconsidering the things at a fundamental level.

A simple fact leading us to say that is, in the case of music (that we can apply as well to other arts) the absolute rupture of the energy continuum between the gesture and the sound phenomenon when we use computer for sound production. Indeed, the computer is a system of symbols, of dynamic symbols, of inter-operating dynamic symbols and we communicate, manipulate, treat these symbols through transducers and interfaces, but in any case we can consider our relation with a computer on the same level than the relation with a physical object. We must consider, in the case of music, that there is no possible assimilation between a musical instrument and a digital sound synthesis system, even in real-time, even with the best possible resemblance between synthesized sounds and real ones, even when the gestural control and every other perceptual aspects of the instrumental relation are restituted. Even in this situation, the computer is not an instrument, but a “representation” of an instrument.

And this is in this role that it plays its more fundamental function, which is a deep revolution.

This is with this major positioning that we introduced the principle of multisensory and interactive simulation of physical objects [8]. We used the computer and its input/output devices in order that all the multisensory-motor conditions corresponding to the primordial instrumental interaction can be **simulated**. This is a kind of iconic representation, in the sense that we try to establish analogy between perceptible aspects of the original and of its representation. But this iconic representation is “integral”, in the sense that the analogy concerns not only one perceptive aspect (shape, colors, etc.), but all the sensory channels, all the action (gestural) channels and the dynamic correspondence between the former and the later during interaction.

So, we introduced the CORDIS-ANIMA language [8] that allows to describe physical objects as an assembly of basic modular components, and to simulate them. Then, we introduced the Force-feedback (TGR) systems [8] allowing manipulating in real-time these virtual objects while we perceive in our hands their mechanical behavior and properties and we hear and/or see their movements, displacements, deformations. Thereafter we developed the GENESIS environment for musical creation with physical modeling, the MIMESIS environment for the animated images and movement creation, and, finally the TELLURIS and ERGOS systems for real-time multisensory simulation [9].

With GENESIS and physical modeling for musical creation, more than ten years of experimentation in laboratory research context and through multiple artistic teaching and creation situations allowed us to create a large "instrumentarium" of physical models for music, and to several composers to create a certain number of musical pieces.

In 2001 [10] we demonstrated that the physical modeling with CORDIS-ANIMA was not only available for the creation of sounds but also for creation process at the level of the macro-temporal musical structure. The basic idea being that physical models of objects that present time-constants on the scale of the gestural behavior can be used, in interaction with physical models of acoustical vibrating objects, to generate sequences of events that can correspond to musical phrases.

We want now present the most recent results and prospects on this way.

4.2. "Supra-Instrumental Interactions"

From the simulation of physical objects at gestural time-constants scale, we deduced several coordinated principles we discuss below.

4.2.1. Gestural time-constant models as simulation of instrumentalists

We call "gestural time-constant models", in the CORDIS-ANIMA representation, the mass-interaction models presenting at least one particle that moves in the frequency range of the gesture, i.e. from 0Hz to about 20Hz. One of the simplest model is an oscillator with, for example, a modal frequency of 1Hz. It can be considered as a (very simplified) simulation of an instrumentalist (a beater), and we can use it to "play", i.e. to hit another (audio frequency tuned) oscillator.

According to this principle, we can build models of complex instrumentalists, and more, of complex instrumentalists interacting between them and with simulations of complex sets of instruments.

Doing that, and looking for adequate parameterization, it is possible to obtain sound sequences of which we can speak, and on which we can work, as musical phrases and musical structures.

In order to avoid here a severe misunderstanding, let's emphasis the fact that it is not question to replace real instrumentalists by simplistic and naive simulations. The stake is much more important and subtle: we have, by this method, a way to model the behaviors and the structure of entities that produces phenomena of the category of the instrumentalist gesture. Then, this becomes a tool to represent, understand, treat, communicate, and teach the gesture. This does not prevent to use it for musical creation, within a physical modeling sound synthesis environment like GENESIS.

The instrumentalist model can be elaborated in order to try to correspond to a real performance, a real gesture, but, conversely, a "gesture" produced initially by a model can be used to inspire new real gestures.

And we can build situations involving real instrumentalists, real instruments, virtual instrumentalists and instruments, in a real-time situation, where these four categories of protagonists can interact (6 ways for interaction are possible between 4 protagonists), in a musical real-time experience or in non-real-time situation where they can together enter in a new kind of creation process.

And finally, retaining only the macro-temporal scale of the phenomena in these types of models, without any obligation of correspondence with the gesture domain, we can develop, consequently, a specific approach of the musical macro structural construction.

4.2.2. The "Supra-Instrumental Gesture"

Let's now introduce something which relates to an enactive approach, no longer at the low multisensory-motor level, but at a completely new scale, and that can be envisaged only in the context of computer and enactive interfaces.

As one can understand, the general framework in which we can work, with mass-interaction physical modeling paradigm, through the CORDIS-ANIMA language associated to TGR (force-feedback devices), allows us to represent complex worlds where real composers, real instrumentalists, virtual instrumentalists, instruments and other kinds of virtual objects, can interact.

To resort to realistic metaphors is a good support to work, and particularly to learn the properties and the potentialities of the system. But one realize very soon that it is yet much more interesting and fertile to escape from the reference to the real world.

A very pertinent situation occurs when we want to produce and control sound phenomena that correspond, may be, to possible natural sounds, but not at all at the scale of the human being. Let's say, for example, the sound of storms, violent winds, sea waves, and telluric catastrophes or, at the opposite scale, sounding microscopic objects.

For such sounds, we have to build models that are generally made of several superposed and interactive layers, at different scales. For example, for a gigantic sea wave, we have to take in charge the sound production at the level of the drop of water, but at the same time, the huge water movement at the scale of the entire waves.

Doing such multiscaled model with GENESIS, we can obtain a convincing result, but we have to imagine a huge gesture, a “supra-gesture” that “plays” the huge amount of water. This doesn’t correspond of course to any possible real gesture and we have to enter in an imaginary poetic universe were we can conceive ourselves as being some giant or mythic personage like Aeolus.

In fact, we have this capability and, even if we don’t have any possible experience of manipulating huge quantity of matter, or blowing huge quantity of water, we are able to conceive such action, as if we had a supra-enactive knowledge, and to conceive physical model doing that.

This, as a lot of various other situations that we could define and describe in these terms, is imaginable and feasible.

To finish, after transposing such scenes in the virtual word of GENESIS, where we have to model large structures (with a very large number of components) that, consequently can be run only in non-real time, we propose the reverse situation. That is to implement on real-time multisensory interactive platforms such models on which we will play by applying a real gesture on a TGR interface.

Then the concept of “Supra-Instrumental gesture” can be understood from these metaphors, and according to its two facets:

1) In a non-real time context, as the “gesture” produced by a large-scale virtual physical system, interacting with a complex and multistructured virtual vibrating system.

2) In a real-time context, where the gesture is real, actually applied through a gestural transducer to a physically consistent virtual object. In this case, the model may correspond to a very large (or very small) physical system (like the previous huge waves) in its structure and in the nature of interaction between its components, but transposed thanks to a kind of “sensory-motor macro or microscope” to the scale of the human manipulation. We can imagine tuning the parameters of the model in order to have the feeling of manipulate a gigantic wave or a nano-object in the hollow of the hand.

5. Conclusion

We introduced, on the base of a quick overflight of important revolutions in the technology for music, the necessity to guaranty an enactive interaction while we are in relation with a computer. This is done, in our research, thanks to the technique of the multisensory interactive simulation of physical objects and the TGR (force-feedback devices) which preserves the ergotic function of the gesture.

Then, we discussed the application of these concepts to the case of the musical creation process through the GENESIS environment and propose an extension of the enactive concept from the normal sized multisensory-motor loops to other scales, thanks to the notion of “Supra-Instrumental interaction” and “Supra-Instrumental Gesture”.

These concepts have been applied for the creation of *Gaea*, a musical piece from the author presented in the Enactive 07 conference concerts and we are now prospecting for implementation of real-time “Supra-Instrumental Gesture” experimentations on TELLURIS and ERGOS platforms.

References

[1] Cadoz, C. “Le geste canal de communication. La communication ‘instrumentale’”, Technique et science informatiques. Volume 13-n°1, pp. 31-61, Hermès, Paris, 1994.

[2] Cadoz C., Wanderley M. “Gesture-Music” – in “Trends in Gestural Control of Music”, M. Wanderley and M. Battier, eds, pp. 71-94 -2000

[3] Kojs, J. “*At and Across*: Physical and Virtual Action-based Music” this volume, Grenoble, 2007.

[4] Goodman, N. “Langages of Art”, Hackett Publishing Comp. Inc., Indianapolis/Cambridge, 1976.

[5] Duchez, M.-E., “Des neumes à la portée – Elaboration et organisation rationnelle de la discontinuité musicale et de sa représentation graphique”, Revue de musique des universités canadiennes, n°4, 1983.

[6] Schaeffer, Pierre. *Traite des Objets Musicaux.* Paris: Seuil, 1966.

[7] Mathews M., Miller J. E., Moore F. R., Pierce J. R. & Risset J. C., “The Technology of Computer Music”, MIT Press, Cambridge (Mass.), London, 1974.

[8] Cadoz C., Luciani A., Florens J. L., “Responsive Input Devices and Sound Synthesis by Simulation of Instrumental Mechanisms : The Cordis System”, Computer Music Journal, 8, n°3, pp. 60-73. M.I.T. Press, Cambridge Mass. 1984.

[9] Florens J. L., Luciani A., Cadoz C., "A real-time workstation for physical model of multi-sensorial and gesturally controlled instruments", ICMC 1998.

[10] Cadoz C., “The Physical Model as Metaphor for Musical Creation “pico..TERA”, a piece entirely generated by physical model”, Proc. of ICMC – Goteborg – 16-20 sept. 2002.

Aknowledgements

This research work has been performed with the support of the French Ministry of Culture and the FP6 NoE Enactive Interfaces.

Topophonies and score navigation in ENIGMES

Roland Cahen
ENSCI, France
cahen@ensci.fr

Abstract

The ENIGMES experimental project took place at ENSCI (National School of Industrial Creation) in 2006-2007. Its aim was to develop usage scenarios of navigable scores or in other words, score instruments. After the end of the project, the aim of this paper is to question the comparison between these experiments as musically oriented topophonies *and the everyday life experience of sound navigation.*

1. Introduction

ENIGMES (Experimentation of New Interfaces for Music And Sound) gathered students, researchers and artists to develop several interaction metaphors: Les Iles, Piège à rêve, Plumage and Feuillage. Each of them exploring different aspects of the navigable score question: small passive haptic, and big graphic physical interfaces, 2 and 3D virtual scores.

2. Interactive scores and navigable instruments

Traditionally, scores are separated from instruments for practical, physical and autonomy reasons. The possibilities offered by digital interactive graphic animated objects, allows instruments to include time reshaping. This means that it is possible to program evolution and interaction into the notation itself. The instrument and the notation can then somehow merge.

3. Sound Navigation

Sound browsing is an everyday life experience. It means moving (as a person or a subject) between different situated sound sources, that are fixed or in motion, interacting with sounding objects or beings. In virtual realities, this can be considered as a kind of realism, unless the aim of the subject's movement is to produce sound and music. Then it can be considered as a way to compose music though motion in space. There are many situations and artistic traditions in which motion and music production are linked together in different ways, such as walking noises and songs, aboriginal songlines, dancing with musical instruments, instrumental music interpretation as percussions…etc. The visual aspect of these music expressions has essentially a space location function.

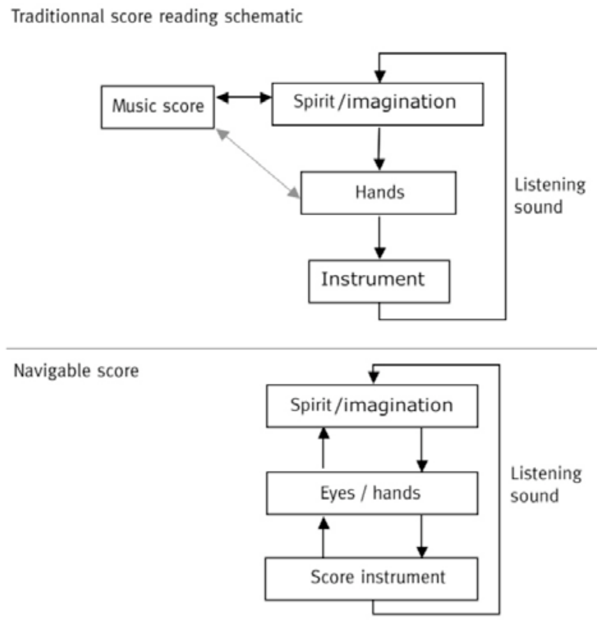


Figure 1. Schematic comparison between traditional and navigable scores

4. Browsing through space and music interpretation

This kind of music *composing while playing* experience is very close to the *enaction* concept in the way it includes a direct encounter with sound-producing objects though movement. The main question is how far musical expression and composition can go using this approach? In which manner can it be designed in order to give users enhanced possibilities to experience music performing? Simply moving in between spatialized sound sources is a very effective and simple interaction modality, but it is limited as a way to perform music, it can be compared with a kind of space mix. But it is possible to implement various other kinds of interactions such as approaching, touching, hitting, and moving objects. It is also possible to modify the object's behavior and use different interaction tools with different proprieties or behavior. We have carried out experiments with some of these interactions in each metaphor.

5. Composing space and time interaction

In the different metaphors we developed, we designed and composed either space or time interaction quite easily, but when trying to do both together, the difficulty was much greater and we only reached rather simple achievements. But the difficulty seems mostly a question of mastering content, method and technique with enough time. Creating *topophonies* requires special skills which are quite unusual together: electroacoustic composition + design and graphics + instrument making + performing.

6. ENIGMES metaphors

We have called our projects metaphors (interaction metaphors), because the easiest way we have to describe and build non-existing object is to use metaphors to make them significant as well a source of inspiration for the design.

6.1. Piège A Rêve

Figure 2. Marco Marini performing on Piège A Rêve

Maurin Donneaud developed for his diploma at ENSCI in 2005 a large textile interaction surface for playing music simply by moving one's hand onto it. During the ENIGMES project, Marco Marini composed an electroacoustic music and extracted its sound corpus for performing. The representation of the music lines were painted onto the fabric. The different music patterns can be played dynamically when touched including time position, control, volume control, triggering, etc. This device is a very effective way to perform electroacoustic music. It is at the same time totally electroacoustic and fully instrumental.

6.2. Les Iles

Figure 3. *Les Iles* interface with concentric circles shapes.

Relief surfaces are fixed on a graphic tablet and are used as passive-haptic and gesture guides for location. Different archetypal shapes are available (concentric circles, merging circles, crossing curves, cushions, rivers, hills, craters, etc.), each of them offering various interaction modes adapted to different sound sets and sound generators. The location within each shape and other parameters coming from the stylus pen are mapped to sound generators in order to offer intuitive play modes.

Les Iles was designed by Marine Rouit and Yiulia Samul.

6.3. Feuillage

Figure 4. A global vision of the interactive score or *Feuillage*

Feuillage is a 2D navigation into a graphic score with different interaction modalities and shape behaviors. We also worked on a principle of collaborative interpretation in which any new interpretation modifies the existing trajectory by reference to statistic criteria. This aspect adds a very interesting point to this kind of application but also makes it more complex.

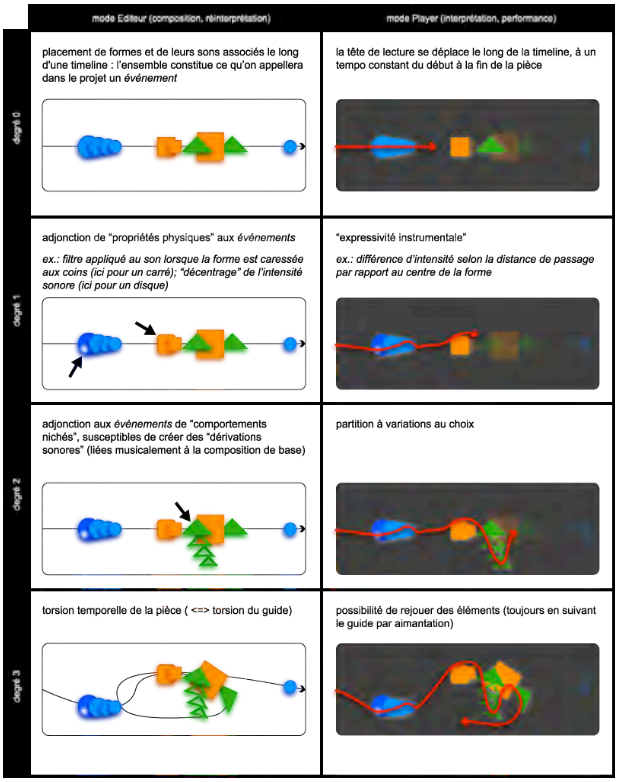


Figure 4. From the canonic composition timeline to a personal interpretation of *Feuillage*

Feuillage was designed by Katalina Quijano and Matthieu Savary with a music creation from Eric Broitmann. Zeev Zohar conceived the collaborative interpretation device.

6.4. Plumage

Plumage is a 3D navigable implementation of Diemo Schwarz's CataRT using Christian Jacquemin's 3D engine VirChor (Virtual Choregrapher). A sound corpus is split into many small grains ordered in the 3D world according to different sound *descriptors*. Yoan Ollivier designed the interaction metaphor referring to feathers; simple shapes giving visibility to each grain with colors, orientation, etc.Three reading heads with three triggers orbiting around each of them play the sounds when meeting the feathers. The point of view and the listening points can also change when switching between different cameras and microphones.

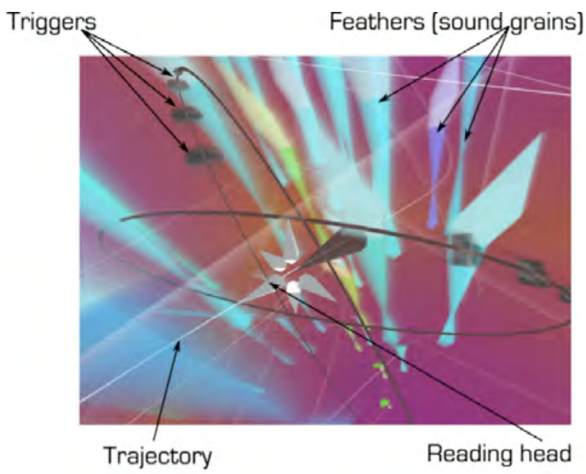


Figure 5. 3D navigation in Plumage

6.4. ENIGMES-Lib

In order to recognize the navigation positions of a pointer into graphical figures for *Piège A Rêve* and *Les Iles*, Jean-Philippe Lambert and I developed a Max/MSP abstractions library (ENIGMES-Lib). It is structured on three different levels:

1) The detectors, which detect positions into rectangles, rings and curves.

2) The mapping section, which scales and connects the position parameters to the generators controls

3) The sound generators or players: *grainplay*, *scratchplay*, *positionplay*, *grooveplay*.

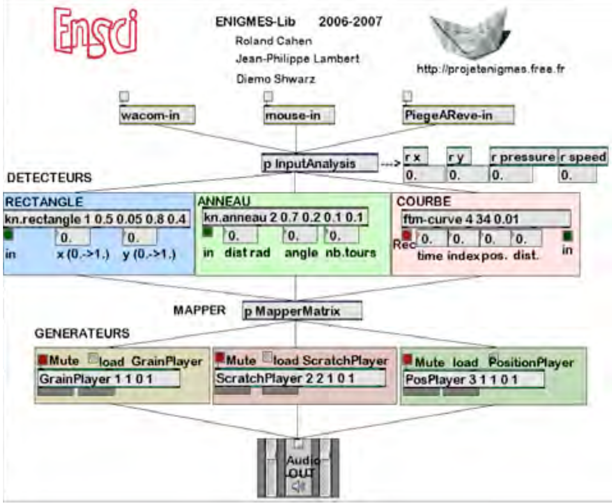


Figure 6. ENIGMES-Lib Max/MSP abstractions

Conclusion

These experiments show that navigable scores can be musical in many different ways than everyday life sound navigation experience and instrumental traditional expression are.

Pour plus d'information

Roland Cahen, ENIGMES La partition Navigable JIM 07

Project website (in French) http://projetenigmes.free.fr Including video presentation and full project report.

http://www.ensci.com (enseignements/ studios de creation/ studio sonore)

References

[1] *Plumage: Design d'une interface 3D pour le parcours d'échantillons sonores granularisés* Christian Jacquemin.

[2] *Visualisation 3D interactive : contrôle et intégration de l'audio* Christian Jacquemin http://sic.sp2mi.univ-poitiers.fr/GT-rendu/articles/Jacquemin.pdf

[3] *Trends in Gestural Control of Music : Gesture – Music :* Claude Cadoz Marcello Wanderley. *Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance* : Joseph Butch Rovan Vincent Hayward. *Mapping Strategies for Musical Performance :* Andy Hunt Ross Kirk (2000, Ircam - Centre Pompidou)

[4] *Concurent constraints models for interactive scores* Allombert, G. Assayag, Desainte-Catherine, C. Rueda http://mediatheque.ircam.fr/articles/textes/Allombert06a

[5] *Musique Architecture (Casterman 1971)* Iannis Xenakis

[6] *Musique et arts plastiques : interactions au XXème siècle (Musique Ouverte MINERVE). Le Son et le Visuel (Dis-Voir)* Jean-Yves Bosseur

[7] *« é c r i t u r e s m u s i c a l e s »* Mémoire de fin d'étude ENSAD Julie Linotte

[8] *"Interfaces Multimodales : Concepts, Modèles et Architectures".* Thèse de Doctorat, Université Paris XI, Orsay, 1995 Y. Bellik

[9] *PHASE Project (Les journées du design sonore2004)* Jean-Philippe Lambert http://mediatheque.ircam.fr/articles/textes/Lambert04a/

[10] *l'Audio-vision (Nathan Université)* Michel Chion

[11] *Music Listening: What is in the Air ? Interactions réflexives : du « c'est marrant » aux machines à Flow* François Pachet

[12] *Sound Navigation in PHASE Installation: Producing Music as Performing a Game Using Haptic Feedback*. Roland Cahen, Xavier Rodet and Jean-Philippe Lambert, Springer Verlag

Lessons from Experienced Gestural Controller Users

Carmine Casciato Marcelo M. Wanderley

Input Devices and Music Interaction Lab
Centre for Interdisciplinary Research in Music, Media, and Technology
McGill University, Montreal, Canada
E-mail: casciato@music.mcgill.ca, wanderley@mcgill.ca

Abstract

Gestural controllers in the musical arts provide a unique perspective on human-computer interaction. Given the novelty of the field, there is little research into the long term usage of gestural controllers. In the course of comparing the Buchla Lightning and the Mathews Radio Baton, the authors questioned users of these interfaces with over 10 years experience. The answers can help performers understand the implications of using gestural controllers, as well as designers who are interested in the long term usage patterns of their interfaces.

1. Introduction

With the rise of inexpensive and powerful computers, sensors, and digital sound synthesis implementations, it is easier to build new digital music instruments (DMI's), and subsequently many have [10] [11]. As a result of all these new designs, the task of evaluating and comparing them has arisen [12] [6]. This is a necessary step if it is hoped that designers and performers will build on the existing corpus of DMI research to help build engaging enactive interfaces for music.

Given the complex relationship between the performer and the instrument [5], it is no wonder that the problem has been approached from a variety of angles. Several visual taxonomies have been presented based on earlier work in Human-Computer Interaction (HCI). Wanderley and Orio [12] presented an overview of these taxonomies, as well as presenting feature sets and evaluation tasks for comparison, and the idea of musical contexts first discussed in Wessel and Wright [13].

In an effort to further the discussion, the authors chose two controllers, the Buchla Lightning II and the Mathews' Radio Baton, for a case study in interface comparison. The entire study entails three parts: a comparison of the technical specifications of each controller, the responses of long term users to a questionnaire regarding their usage patterns, and the motion capture of performances using these controllers. This paper details results from the first two parts.

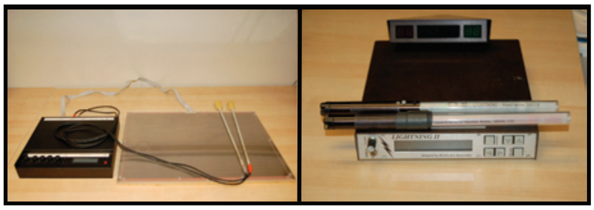


Figure 1: Radio Baton (left) and Lightning II (right)

2. History of the Controllers

The following section gives a background on the controllers to better situate them in the user responses. Figure 1 shows both controllers. Figure 2 shows a modified dimension space [1] so as to quickly highlight the major technical differences.

2.1. Buchla Lightning II

The Lightning was developed in the late 1980s along with three other prototypes called Wind, Rain, and Thunder: of these, it was the only one to go into production. The prototypes were a response to what Buchla felt was a lack of original MIDI controllers on the market [4]. The Lightning II triangulates the infrared signals from handheld wireless 'wands' providing a large 2D (x and z) area for the user to freely gesticulate within. It is going through its third revision and is marketed through Buchla and Associates [1]

2.2. Mathews Radio Baton

The Radio Baton is the culmination of a number of drum-inspired controllers by Mathews, including the Daton and the Sequential Drum [8] [9]. The Radio Baton was created using a variant of capacitive sensing developed by Bob Boie at AT&T Labs in the late 1980s.

[1] www.buchla.com/lightning

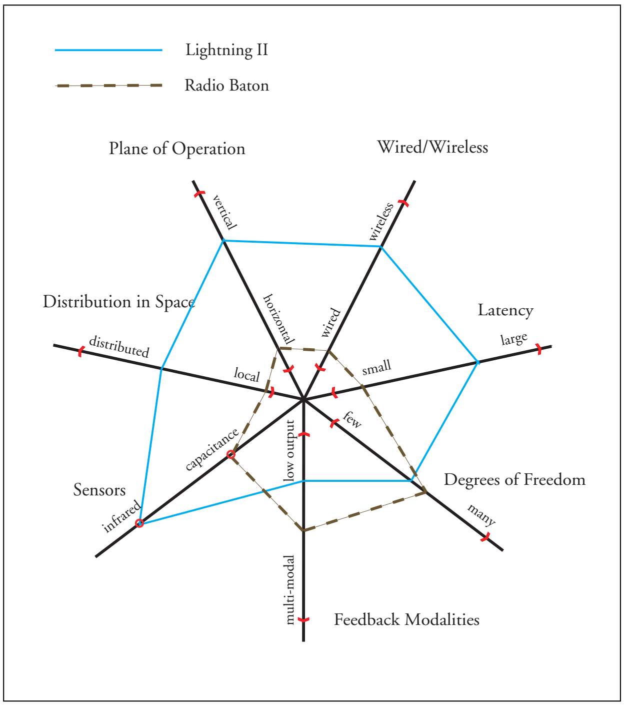


Figure 2: Radio Baton and Lightning II plotted on a dimension space [1] so as to clearly contrast their technical characteristics.

Using five electrodes over the surface of a square plate, two wired antennas that resemble batons would increase the capacitance value when near an electrode. The differences between electrode capacitive values are used to deduce the xyz position of each baton [2]. It has been used in over 40 pieces [3] and is marketed by MarMax[2].

3. Long Term User Responses

In order to get a sense of usage over time, questionnaires were sent out to the known professional users of each controller, whiched totaled under ten. Five questionnaires were received, three from Lightning users and two from Radio Baton users. This small set of responses is to be expected when so few are known to play the instruments under study. The questions and resulting answers were informal in nature thus contributing an important qualitative aspect to the present study. Each question and response is reported in the following tables. Answers have been paraphrased for presentation. Performers 1-5 correspond respectively to the three Lightning users and the two Radio Baton users.

The amount of experience amongst the respondents shown in **Table 1** is noteworthy. It has been proposed that over 10 years of formal training on a particular acoustic instrument qualifies a musician as an expert [7]. While there is no formal pedagogy for gestural controllers, the ability to become an expert user of these devices given time should be assumed, in the same way as race car drivers are assumed to be more skilled drivers than the average.

[2]223 Precita Ave, San Francisco, California, 94110

Table 1: **Experience**

How long have you been using the Lightning/Radio Baton?

1	14 years
2	15 years
3	15 years
4	20 years
5	20 years

Table 2: **Usage**

What were some of your favorite uses of the Lightning/Radio Baton?

1	solo performance, accompaniment of dancers
2	performance of compositions for Lightning, accompaniment of silent films, children's theatre, medical rehabilitation
3	solo performance, as part of jazz ensemble, score-following, conducting, spatialization
4	performance of compositions for Radio Baton w/ voice, orchestra, manipulation of video, spatialization
5	performance of compositions for Radio Baton, improvisation with acoustic instrument musicians, as part of world music ensemble

The responses to the question in **Table 2** seem to indicate that the controllers are extremely adaptable to as many different music performance settings as established acoustic instruments are, (and some that acoustic instruments cannot be used in, such as the manipulation of video media). Also evident in the answers is that many opportunities to perform arose almost solely from the novelty of the controllers, the "shiny new gadget" phenomenon. For example, Performer 3 was approached to perform for millenium celebrations in Times Square, New York for what he supposes was this reason. The novelty aspect also had drawbacks: Performer 1 recalls a dance accompaniment performance where the program notes did not specify his usage of the Lightning. As a result, the audience perceived the movements as dance-oriented rather than those of a music instrument performance. Designers would do well to use the novelty of their designs as a means to reach potential performers while being aware that audiences are not as comfortable with new interfaces as with traditional instrumentation.

Almost all respondents are either percussionists, pianists, and/or conductors and admit their approach to these percussion-inspired interfaces owes a great deal to their acoustic instrument training. Performer 2 finds that

Table 3: **Musical Education**
Do you play any other instruments, acoustic or electric? How has this informed your performance/compositional technique regarding the Lightning/Radio Baton?

1	acoustic mallet percussion, movement training in theatre/circus skills
2	piano, percussion
3	piano, conducting
4	guitar, trumpet, piano, voice, dataglove, theremin, laptop
5	percussion

"realizing when (and when NOT) the gestural repertoire of the Lightning fits percussive technique and sounds sent me off on a path of exploration that continues today". Performer 5's percussion training has raised his expectations of the performance behaviour of the Radio Baton in terms of strike latencies. Performer 1 recognizes that the emphasis in mallet percussion on proper mallet placement in space has influenced his approach to the Lighting. He also mentions the theatrical nature of playing the Lightning. Finally, Performer 4 notes that musicality can be transferred: "the more musical you are, the more musical your music and your performance of ANY instrument". Contrasting these answers with those from **Table 2** suggests acoustic musicians can take new interfaces beyond their usual performance settings, all the while drawing on their acoustic instrument training.

Table 4: **Role**
What role have you played in the development of the Lightning/Radio Baton?

1	Lightning I repairs, primary technician for beta-testing, preset development, soldering and assembly of Lightning II, circuit board layout for Lightning III
2	none
3	none
4	composer/collaborator
5	none

Regarding the answers in **Table 4**, although only Performer 1 explicitly worked on the development of these interfaces, all respondents had close relationships with the controller designers. Performer 2 eventually collaborated with Don Buchla to build other controllers and Performer 3 had several conversations with Buchla about possible design improvements to the Lightning. Performer 4 has worked closely with Max Matthews; his "musical use of the system influenced the design of the hardware and the features of the software." It is interesting to ask whether being in close contact with the designer kept the musicians motivated to continue using the interfaces.

Table 5: **Approach**
Do you approach the Lightning/Radio Baton as an instrument or a controller?

1	both
2	instrument
3	gestural controller
4	instrument
5	both

The answers to this question can be viewed from two perspectives. The first revolves around the nature of a gestural controller in the schema of a DMI. In Performer 2's own words:

> "A controller is a device that transduces a physical gesture to a defined control signal. The signal is then sent to some sound-producing apparatus. A musical instrument is...a device or system...for real time musical performance... concerned with expressive and reliably repeatable control of...musical parameters."

In this sense, Performers 2 and 3 utilize the sound engine of the Lightning with its mapping abilities in many of their performances and therefore use it as an instrument. It is impossible to do so with the Radio Baton as it does not have it's own sound engine, thereby rendering it a gestural controller by default. Performer 4 substantiates this by mentioning "the Radio Baton is what turns my laptop into a musical instrument."

The second perspective is more philosophical in nature involving the question of when does a gestural controller become a musical instrument. Performer 1 feels the Lightning is "a controller that only becomes an instrument after a lot of practice". This is echoed by Performer 2's remark that "almost anything can be an instrument in the hands of an accomplished player". This would seem to support the claims regarding expertise in [7] in which time spent practicing is essential. However, Performer 5 approaches these terms as states of operation which can be intermingled in a performance: "I use it in both modes, sometimes simultaneously and always jumping from one mode to the other. To me it would not be worth playing if it couldn't occupy both roles."

For Performer 2, each different mapping scheme creates a new instrument. As evidenced in the responses in **Table 6**, these interfaces can take the (conceptual) form of many existing and novel control paradigms which do indeed resemble completely different performance practices. From Performer 4 "It is NOT a percussion controller - that limits its function, subtlety, and role to something too primitive and simple. The construction

Table 6: **Mapping**
What have been some of the more successful mapping strategies you have employed with the Lightning/Radio Baton? Do you find certain paradigms or mapping strategies more easily 'fit', for example as a percussion controller or "an expressive tape player"?

1	virtual conductor, mallet keyboard, large drums, diatonic scales played with horizontal movement and button switch, use and tuning of sound engine presets
2	percussion controller, control of continous parameters spatially, piano controller, layers, use and tuning of sound engine presets
3	virtual conductor, max patches and external synths
4	conductor, soloist, improvisor, timbre sculpting, note-based, remixing, spatialization, triggering
5	flying over surface to trigger events, image/video controller

does influence its perception, but not it's use or potential."

Table 7: **Modifications**
Is there anything you would change (or have changed) about the Lightning/Radio Baton in terms of ergonomics, technical performance, etc.?

1	**Would change**: strike latency time of 40 ms
2	**Would change**: software for editing presets, wireless connection between head and sound engine, haptic feedback in wands
3	**Would change**: sensing of wand position in all 3 dimensions, more ergonomic wand grip, change wand to one resembling conductor baton
4	**Have changed**: software has allowed for all necessary changes
5	**Have changed**: foam moved to surface, batons changed to modified drum sticks, computer interface through audio interface

The changes proposed by the Lightning users shown in **Table 7** run the gamut between performance, usability, and ergonomic improvements. Performer 5 has improved the Radio Baton such that he satisfied with the performance of it. His main problem now is just to "create sounds that are vivid and malleable". Performer 2 has invented his own notation for scoring his Lightning performances, as well as a symbolic notation for modifying and creating presets on paper.

4. Conclusions

Several interesting points seem to emerge from the responses. It is evident that gestural controllers can be designed well enough to sustain the interest of musicians. Given access to the designer, the musician will suggest design changes or even undertake the changes themselves, thereby prolonging the life of the controller.

Also, there seems to be a correlation between the physical interface of these controllers and the acoustic musicians it attracted. As the sample for this questionnaire is much too small to make any large generalizations, it is still noteworthy that percussionists and conductors were drawn to baton-based interfaces. Also, this did not limit the performance contexts in which they were used. Respondents included many different types of performance contexts including video and spatialization control.

Mapping is at at least as important to musicians as the physical interface, and even more so over the long term. Using a different mapping strategy results in a new control paradigm to explore. This ability to change such an elemental part of the instrument seems to be part of the appeal of this type of musical interface.

References

[1] Birnbaum, B., R. Fiebrink, J. Malloch and M. M. Wanderley "Towards a dimension space for musical devices" in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2004.

[2] Boulanger, R. "The 1997 Mathews Radio-Baton and improvisation modes", in *Proceedings of the International Computer Music Conference*, 1997, pp.395–398.

[3] Bresin, R. , K. Falkenberg and S. Dahl. "The Radio Baton as configurable musical instrument and controller" in *Proceedings of the Stockholm Music Acoustics Conference*, 2003, pp.689–691.

[4] Buchla, D. Keynote Speech, *Conference on New Interfaces for Musical Expression*, 2005.

[5] Hunt, A. and R. Kirk "Mapping strategies for musical performance", in *Trends in Gestural Control of Music*, eds. M. Wanderley and M. Battier, IRCAM, Paris, 2000, pp. 231–258.

[6] Jorda, S. "Digital instruments and players: Part I - Efficiency and Apprenticeship" in *Proceedings of the Conference on New Interfaces for Musical Expression*, 2004, pp.59–63.

[7] Lehmann, A. C. "The acquisition of expertise in music: efficiency of deliberate practice as a moderating variable in accounting for sub-expert performance" in *Perception and Cognition of Music*, eds. I. Deliege and J. A. Sloboda, Psychology Press, 1997, pp.161–187.

[8] Mathews, M. *The Sequential Drum*, IRCAM, 1980.

[9] Mathews, M. Worskop Presentation, *Conference on New Interfaces for Musical Expression*, 2003.

[10] Miranda, E. and M. Wanderley *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard.* A-R Editions,2006.

[11] Various Authors *Proceedings of the Conference on New Interfaces for Musical Expression* 2000-2006.

[12] Wanderley, M. and N. Orio. "Evaluation of input devices for musical expression: Borrowing tools from HCI" in *Computer Music Journal* 26(3), 2002, pp.62–76.

[13] Wessel, D. and M. Wright. "Problems and prospects for intimate musical control of computers" in *Computer Music Journal* 26(3), 2002, pp.11–22.

Computer-Mediated / Interactive Performance: *Moving Boundary Problem* Case Study

Thomas Ciufo[*] David Birchfield[*†]
(*)Smith College, USA
(†)Arizona State University, USA*
tciufo@smith.edu, dbirchfield@asu.edu

Abstract

Moving Boundary Problem is a multi-channel, interactive live electronic work for two performers. The piece has developed around a gestural language that explores the sonic and expressive capabilities of a pair of hybrid acoustic / electronic instruments. Utilizing a wide range of human / computer interfaces and signal processing techniques, these composed instruments extend acoustic sound sources including found objects, flutes, and percussions. *Moving Boundary Problem* is a manifestation of the unique acoustic, gestural, and human relationships that emerge from interaction with and through these new instruments. In this paper, we discuss the conceptual, aesthetic, and technical concerns encountered in the ongoing realization of this improvisational live performance project.

1. Introduction

Improvisation is a broad and complex topic [1, 2, 3] that places unique demands on its practitioners and the tools they use. As composer / improviser Pauline Oliveros points out, “The improvising musician has to let go of each moment and also simultaneously understand the implications of any moment of the music in progress as it emerges into being” [4]. This type of in-the-moment creation requires a flexible / pliable instrument, and a deep intimacy with the instrument’s sonic potential. This is the case whether the instrument is acoustic, electronic, computational, (or some combination), with each type offering unique challenges and possibilities. A number of practitioners have explored a comprehensive approach to improvisation that spans across instrument design, composition, and performance. Bahn and Trueman provide a definition of composed instruments, explaining, “We consider our entire systems, from physical instruments, sensor interfaces, interactive computer music environments to spherical speaker arrays, to be both extended instruments and non-linear compositions: composed instruments” [5]. Commenting on the use of computational frameworks in improvisation, Dean points out, “One of the powers of computer-generated sound is to progressively move along axes that distinguish one sound from another by a variety of morphing processes. Thus one can choose to use a sound world aligned to instrumental sound, to natural sound, or to electronic sound, and to move between such worlds freely, and even gradually” [6]. One of the most enabling aspects of these computer-mediated performance systems is the real-time access to a complex and malleable sound world that is not restricted to the domain of acoustic instruments, but rather combines the processing and transformative power of the computer with the richness and immediacy of the acoustic world.

Our work in *Moving Boundary Problem* draws on related research in the design and realization of composed instruments, and considers the formidable challenges of improvisation in this context. We are designing performance systems that have a range of sonic possibilities that can be navigated or explored in the context of improvisational performance. Playing these extended instruments in a duo context poses new challenges, which then inform the design and refinement of the systems and our performance approach. By continuing to play together over time, both the instruments and the improvisers change and adapt, responding to idiosyncrasies that make us (and our instruments) unique.

In the remainder of this paper we will discuss our individual performance systems, and then articulate the ways in which we use these systems to explore / enact interactive sonic performances.

2. Ciufo: Instrument Design

The performance system that I am currently using has grown out of a lineage of composed instruments designed for live improvisational performance [7, 8]. The overarching system design includes specialized hardware, custom software, and a range of tactile sounding objects that form an integrated live performance system. One important design criteria for this system is a high degree of pliability, which makes it usable in a variety of performance contexts, including solo or ensemble improvisations.

Figure 1. Ciufo in performance

This system uses a microphone and contact mic as the main input devices. This allows me to bring into this system a wide range of sound sources, including found objects, small hand percussion instruments, gongs, bells, and flutes. The hardware design consists of a computer, audio interface, and physical control interfaces, including both switch and continuous control pedals, and a fader based control surface. The output sound is spatialized into a 2D space using either four or eight speakers. The software design, (realized in Max/MSP/Jitter [9]) has evolved over time, and has been described elsewhere [10]. The initial use of this composed instrument was in the context of solo improvisational sound explorations, enacted over time frames ranging from ten minutes to nearly an hour. For *Moving Boundary Problem*, I am expanding and refining my use of this system in a collaborative duo context, and learning how to engage the improvisational approach of another musician.

A flexible signal processing architecture affords a wide range of available sound transformations. Thus, instead of building the sound processing network using a fixed signal flow, this system uses a modular matrix mixing / routing design. All signal processing modules are connected to a two-dimensional signal patching matrix that can route any input to any output at any level. This enables dynamic and continuously variable signal routing and mixing, including serial, parallel, and tree structures, with feedback and adjustable delay times available at each node. An example signal path is shown in Figure 2.

Available signal processing modules include timbre / frequency shifters, signal decimation and clipping, variable rate time stretchers, delay modules, a tunable comb filter, a noise-based resynthesis module, a stereo granulator, and a number of interactive recorder / player modules. This modular design also supports changing out specific processing modules, and reconfiguring the signal flow using simple connection templates. In performance, these templates can be recalled with a single continuous controller used to interpolate between two states, allowing for complex transformations using simple, metalevel controls.

Because of the reconfigurability, complexity, and high number of controllable parameters available with this system, developing and internalizing interaction strategies has been quite challenging. As discussed elsewhere [7] the relationship between autonomous, manual, automated, and analysis-based controls will greatly influence the behavior of a performance system, with certain relationships proving more appropriate to particular types of interaction. While this software is still capable of unexpected results, for *Moving Boundary Problem*, I have scaled-back some of the autonomous system behaviors, choosing instead to focus on the human-to-human interaction that so richly informs our duo improvisations.

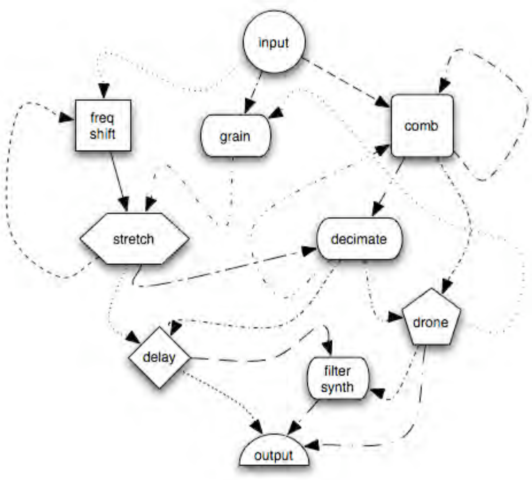


Figure 2. One possible Ciufo signal path

3. Birchfield: Instrument Design

Figure 3 shows a hybrid percussion/laptop performance system designed by Birchfield. Central to its design are the concepts of extended and composed instruments. In this instrument I have sought to leverage the expressive attributes of conventional percussion instruments and to reinforce the diverse timbral possibilities of a collection of skins, metals, and woods. The instrument is comprised of a fixed set of source sounds including found objects, folk instruments, and conventional percussion sounds. These include: electronic drum triggers, a large shaker, a block of zebra wood, a suspended Chinese cymbal, a large Indian bell, and a kalimba. Each of the acoustic

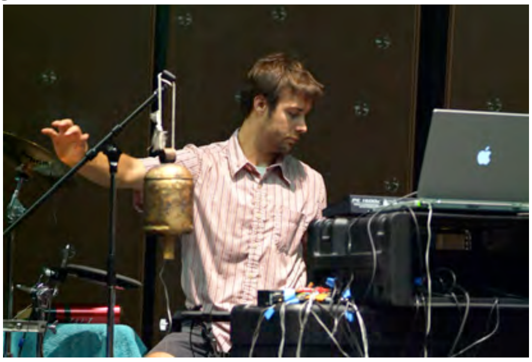


Figure 3. David Birchfield in performance

instruments has a piezo-electric contact microphone that is discretely routed to a multi-channel audio interface and into an interactive software framework that is written in Max/MSP. As pictured in the Figure 4, each sound source is manipulated as a separate signal processing chain that is tuned to an appropriate set of processing techniques and parametric controls for each instrument. A set of data control foot pedals and physical slider banks provide the performer with direct control of software synthesis parameters. Multi-

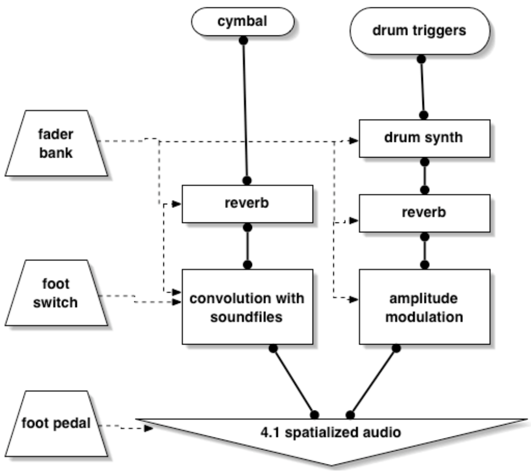


Figure 4. Birchfield instrument architecture

channel, spatialized audio is routed out of the software.

This instrument is designed to engage several aspects of conventional percussion performance. First, the distributed arrangement of discrete instruments is derivative of a conventional drumkit. This physical arrangement supports full body striking motions that coordinate between both hands and feet of a seated musician. Just as in conventional percussion setups, the performer has full access to a large configuration of instruments that can be selectively utilized.

Second, the instrument design seeks to leverage the nuances of physical interaction with these instruments. For example, the Chinese cymbal can be excited by a number of sticks, mallets, or the hands through a variety of physical gestures. Each gesture will yield different spectral shapes and amplitude envelopes. These subtle shifts in physical performance are further colored by the use of flexible signal processing chains in the software. Third, the instrument extends the acoustic sound production paradigm such that if the performer does not provide sound input, the software will not generate sound output. This places control and responsibility on the performer in a manner that is familiar for a percussionist.

In prior work, I have implemented percussion based performance systems that were intended for solo performance [11]. While this prior instrument yielded successful results, it was not well suited for collaborative performance. In contrast, this current instrument is designed for ensemble interactions. It provides a malleable interface to sound that is rooted in physical gesture. Thus, the instrument can be adapted to match or contrast with other musicians and their instruments, and its dependency on physical gesture facilitates human-to-human interaction that is critical for ensemble performance situations.

4. Player Interaction

We have described the factors and considerations that underpin the realization of these two extended instruments, and have discussed how these sound interfaces lead to idiosyncratic outcomes that reflect our individual aesthetic and design decisions. Here we describe our process of improvisation and articulate how our collaboration leads to distinct musical outcomes that are rooted in both sonic and physical relationships.

Despite differences in their basic architectures, our instruments are sonically related through overlaps of acoustic sound sources and signal processing techniques. For example, both systems utilize an array of percussion instruments including bells, gongs, bowls, and cymbals. Similar sonic overlaps arise as Birchfield draws on a large database of collected soundfiles that include recordings of bird songs and the voice. These can be recalled to draw associative relationships with Ciufo's acoustic bird call and flutes. Finally, both performance systems utilize a combination of pitched and non-pitched sounds. Ciufo plays a collection of wind instruments while Birchfield employs a bell and kalimba to spawn pitched material.

The use of similar digital signal processing techniques further contributes to our shared sonic language. Specifically, both instruments use granular synthesis, convolution, and delay loops to process acoustic and digital sound sources. The application of shared synthesis techniques that can be applied across a body of related sound sources provides rich opportunities to expand into new sonic idioms that are specific to our collaboration.

Our instruments are physically related through a body of overlapping performance gestures and transparent interfaces to sound. This physicality enriches our human-to-human communication during performance. For example, each instrument includes percussive instruments that can be played in similar ways. Birchfield's instrument features a suspended Chinese cymbal that can be played with the hands, mallets, or sticks with a variety of gestures including striking, sweeping, rubbing, sliding, glancing blows, or rolling. While playing bells, bowls, and gongs, Ciufo uses a similar repertoire of gestures with sticks, hands, or a bow. These physical interfaces communicate a great deal about upcoming sonic events. In many instances we will perform the same physical gesture that may lead to divergent sonic outcomes. For example, sounding a temple bowl with a circular rubbing motion in Ciufo's instrument will sound different from a circular sweeping motion on the cymbal of Birchfield's instrument. Nonetheless, these types of shared physical gestures often serve as important structural moments in our music.

5. Outcomes and Conclusions

How does this combination of individual instruments, unique personal aesthetics, and shared player interaction coalesce into a meaningful outcome? How are decisions / choices made in the context of improvisational performance, and what constitutes meaning within this context? In primarily idiomatic improvisation, there may be a variety of formal / contextual frameworks that influence the musicians' choices. Even so, many musicians are not consciously aware of how they improvise. Doc Cheatham confesses, "I have no idea of what I am going to do when I take a solo" [12]. This scenario becomes even more complex in improvisational environments without defined idiomatic boundaries. Even in non-idiomatic, or so called free improvisation, the players bring a range of influences, biases, and preconceived notions to their playing. Additional biases or predispositions are built into most composed instruments.

In *Moving Boundary Problem*, we explore how our individual sound identities and performance aesthetics interact and coalesce into unique and expressive sonic outcomes. The previous section describes aspects of our moment-to-moment interaction, focusing on both physical and sonic gestural relationships. Our work is also built upon a commitment to active, careful listening, and a willingness to follow one another down interesting sonic pathways. This is often a type of associative interaction, in which a particular sound will suggest a certain response, possibly supportive, complimentary, or contradictory.

We often provide complementary responses or extensions to each other's sounds or gestures by leveraging the capabilities of our individual instruments. For example, Birchfield's frenetic full body rhythmic activity may be supported by Ciufo's shifting textural underpinnings. At other times, we may both explore a similar set of sounds, such as bowed or scraped metals, thus blurring the line between the individual performers / instruments and the resultant aggregate sound. Our focus on the manipulation of sounding objects supports dynamic and synchronized interaction that are only possible in the presence of physical gestures.

This back and forth / mutual influence, combined with a sensitivity to overarching development and formal cohesion informs much of our work together. At the same time, there remains an aspect of our work that is indefinable, and even a bit mysterious, occasionally producing totally unexpected results. This seems fitting for a practice that Bailey describes as "the most widely practiced of all musical activities and the least acknowledged and understood" [1].

References

[1] Bailey, D. *Improvisation: Its Nature and Practice in Music*. Da Capo Press, 1993.

[2] Nettl, B., and M. Russell, Eds. *In the Course of Performance: Studies in the World of Musical Improvisation*. Chicago Studies in Ethnomusicology. University of Chicago Press, 1998.

[3] Dean, R. T. and H. Smith. *Improvisation, Hypermedia, and the Arts since 1945*. Harwood Academic Publishers, 1997.

[4] Oliveros, P. "Quantum Improvisation: The Cybernetic Presence." *Improvisation Across Borders Conference*. http://www.deeplistening.org/pauline/writings/quantum.html, 1999.

[5] Bahn, C. and D. Trueman. "Interface: Electronic Chamber Ensemble." *New Interfaces for Musical Expression at the CHI2001 Conference for Human/Computer Interaction*. Seattle. http://www.arts.rpi.edu/crb/interface/interface-Bahn and Trueman.pdf, 2001.

[6] Dean, R. T. *Hyperimprovisation: Computer-Interactive Sound Improvisation*. A-R Editions, 2003.

[7] Ciufo, T. "Design Concepts and Control Strategies for Interactive Improvisational Music Systems". In *Proceedings of the MAXIS International Festival / Symposium of Sound and Experimental Music*. Leeds, UK, 2003.

[8] Ciufo, T. "Three Meditations for Prepared Piano and Computer". http://www.ciufo.org/research.html, 2002.

[9] Cycling'74. Max/msp/jitter graphical audio and video environment (computer program). http://www.cycling74.com/

[10] Ciufo, T. "Beginner's Mind: an Environment for Sonic Improvisation". In *Proceedings of the International Computer Music Association*. Barcelona, Spain, 2005.

[11] D. Birchfield, "Composing the Digital Rainstick". In *Proceedings ACM Multimedia*, New York, NY, 2004.

[12] Berliner, P. *Thinking in Jazz: The Infinite Art of Improvisation*. University of Chicago Press, 1994.

The instrumental space

Frédéric Curien
Ecole Européenne Supérieure de l'image EESI, France
f.curien@eesati.fr

Abstract

A musical instrument produces sounds; it is a sonorous element, a sounding body. It is a necessary but not sufficient condition. It is not the prolongation of the body, but in the prolongation of the ear and the intention. It is a relation with an intention, which is neither in the imitation, nor in the reproduction. This relational space which goes far beyond the instrument as such can be qualified as instrumental. It is composed of the entirety of the bodies which are in an instrumental and enactive relationship, the body of the object - tool, that of the interpreter, the body of the texts, and the social body.

1. Introduction

A tape recorder is not a musical instrument but a device of sound reproduction. In order for it to become one, one must put it in relation to a musical intention, transform it into an instrument, create accesses, and implement a musical and technical structure. Similarly, a microphone is a device whose unique function is to capture sound and transmit it into in the form of a standardized electric signal to other devices which will reproduce it with maximum efficiency, by respecting the best signal/sound relationship. If one uses the microphone as a device, in its specific functionality, then one can consider whether it satisfies the requirements of one's schedule of conditions or not. For the microphone is not transparent to the matter that it captures. It undergoes transformations by reason of the distance of the captor in relation to the source, the position within space, the tone varies according to various parameters. The person who masters these shades, is capable of using them, making them vary, of reproducing them and thus influence the sound result with sufficient prediction, so one could say that he is forming an instrument with the device. Transforming the device into a musical instrument entails approaching the object, and being at once with it, using it with an idea, and "when a thing is perceived through an awareness of temporality, it is turned into something which is no longer a thing. This feeling of entire (engulfment) gives a mental depth to the object so that it ceases to be a simple object and becomes art.

So, any sounding body used by the composer can become a musical instrument, for Berlioz in his Big Treaty of modern instrumentation and orchestration in 1843, almost as Marcel Duchamp later with its ready-made. The instrument is an image of society, a relation in which the adaptability is a major characteristic. An instrument fits into a set, a system which has no "project". All the concerned "bodies" have their definitions of the instrument, according to their manners and to their reference systems. The set of these definitions brought to the foreground of the multidimensional notion of the instrument, so much as the contact areas. Its "physical" membranes are multiple. The instrumental challenge lies in the contact points between the reality which resists, foundation of the sensation, and the virtual (model, number, discreet): as Edmond Couchot underlines it, it is for the interface of reality and the virtual that the artist is called to rehouse. « Narrow but fertile margin where glance and calculation interpenetrate, the extreme touches extremely ». It is at stake, because in this limit witch is always more infra thin, is situated the pleasure of the game (to blow, pull, to manipulate) and the discovery with the body in the space of the other.

Figure 1. The practice is connected to different types of memory (auditory, visual and digital) to which is added a memory which is not only of a postural nature but also muscular and proprio perceptive.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

2. An instrument is a network

A musical instrument is not a device or prosthesis, because it does not fill a lack, an absent organ. It is not either an orthesis which would raise a deficient organ. An instrument is at once an object, which one holds close to one by its nature, its material and the invoice, and it is also the tool, which serves to create a current of communication, which opens a spatiotemporal canal with the other. The instrument is not a continuation of the man, but rather an interface. The characteristic stand in a relation, an adaptative answer in the middle, not evolutionary in the biologic sense, because he does not raise the genetic from the transmission of acquired, constructed characters. It is not passed on by heredity, but by inheritance. But, it includes a code, the registration of a musical system. It is passed on, with the series of the gestures which shaped it, from generation to generation. As the tool, the musical instrument is of the order of culture (of the rule), and it is only as a relation that it becomes integrated into art (in the exception). It is free as an object of craft, to go into the artistic sphere. It is thus the inextricable character of these two aspects, manufactured object and tool of culture which makes its specificity. One can thus, by extension, admit to this reflexion all the surrounding actors, all the instruments of thought, and the surfaces of registration of the thought and the gesture (language, word, concept, partition) which are of the order of tools, and return in touch with the space.

The forms that set musical instruments and way of using it vary historically. It exceeds the frame of this article to make a description of it but the adaptability of the instrument in its destination and in the preservation of the current events, is a condition of the survival. It is necessary to take here the term "evolution" in the sense of change and adaptation, not to consider it according to the angle of the progress, suspected of positivism. Technological performance is real locally, but the idea that the instrument would be perfectible in the infinity recovers rather from an ideological position (perfectibility in the infinity: metaphor of the unlimited growth). In the era of virtual technological, the instrument continues to be for the musician a crucial point of search where requirement confronts in times of answer of the order of the millisecond. In the musical world, there is a real technical culture, because there is a tradition of the instrumental investigation. Musicians wonder how a computer works as they sound a violin to take out its sound. If a musician needs a "real" piano, it is to use it in all these dimensions of sounding body. The computer is a tool because it allows to make, in a poietic sense, exactly because it is not a guitar, it is not a bassoon, but it can be the other thing (matter) for the artist, because his essence is that of a mathematical being.

3. An instrument is "controllable"

Devices cut the space and the time in a way which is appropriate for them. The cut, streaked space, is that of the device and the fluid and continuous smooth space, that of the body. So it is not the machine which inaugurates this discretisation in the industrial era, because the musical instrument cuts the space in edges, in holes, according to a register and modalities which form ambitus, in degrees and in ranges, a system of scales which implies series, as any language goes discretised world to units manipulable and opposable. It is not the machine and even less the computer which inaugurates this spatial discontinuity. Difference is in the change of scale and global socioeconomic context. The machine creates another relation to man with the material (subject) where the individual is only a link. Certainly, without the mastery of the man, no technical complex could work. Since André Leroi-Gourhan, one does not set any more language and tools which are the expression of the same human property, “for both, there is a reality and an appearance of their mastery by the user, the appearance and the reality of their mastery on the individual before whom they preexist and who, without them, would not be”.

A musician has a report in the mastered object, its instrument: technical mastery is a criterion of professionalism. Music preserved a tradition of mastery. For the composers, it is a value which guarantees the life of the written work, and which is estimated: mastery of the gesture, the reading which is also an instrument, of listening, the coordination of the group of the body. The composer is as the architect who makes his plans on paper, but has to respect the techniques of construction. Without this system, score does not exist, will not be restored. Moreover Max Mathews, pioneer in computer music, compares more easily the computer with an orchestra, with a group of instruments, rather than with a solo instrument. The composer takes with the computer the function of the leader, which steers a technical set and he is responsible for the coordination of his group. For the artist using the computer, to deny the interest of a technical mastery of its tools which join by their conception and functioning a technico-aesthetic set, well engage him in spite of him ideologically and aesthetically. From 1961, the minimalists refuse the instrument to find objectivity “What you see is what you see”. They look for materials which do not offer harshness, defects which could be seen, neutral materials on which glance slides. The body of the artist is not evacuated, it remains in the process, in the studio, but it is the body of the other whom brings in a game, the artist goes out of the game, the subjectivity, but the space, the place and the body of the spectator form the device. The technique is not craft any more, which shows the gesture (the track of the body, the gesture), but becomes industrial: neutrality of the building material; to see Tony Smith or Robert Morris with the manufactured object, without a track. The artist is a designer; his idea takes him on the effectuation which has no more right of citizenship. There was a loss of the instrument, constituted traditionally with the paintbrush the canvas and the building material, for a vision widened with the instrument in the device.

Visual art in the evolution towards the immateriality, the disembodiment, cuts itself of the instrument, by returning in break with the technical tradition, well before the arrival of the computer. He does not ask any more the question of the conservation of works and claims the short-lived (to see performances, installations or procedural works).

4. Ideas and conception, new perspectives?

New paradigms appear which do not only arise from the technique but from the social organization in its completeness. Environment changes as well as way of life, way of living and working customs, networks of exchange and communication develop. There are at the same moment technological and cultural transfers with a bilateral synergy. As regards the art, it is interesting to consider the example of music where exchanges among the composers and their instruments are made through a corporate body, the factors of instruments or stringed-instrument makers. The need of an instrumental innovation often arises from a need engendered in the expression of the idea, such as it existed at numerous periods in history of music and organology. The principle of the “double escapement” of Sébastien Érard, patented in 1821 and worked out for Liszt is the example of it the most known. It is interesting also to consider the other cases where technical feasibility is at stake in front of the urge to realize a work. The example of Piotr Kowalski and his technological utopia, his dream of technique is very telling: the use of techno sciences is not the end in itself but allows to know reality for this artist: " the more one has of knowledge of the reality, the more the imaginary has of things to be handled " it is necessary to open not in new technologies but in questions which these technologies allow to ask.

5. Episteme

Episteme: the model as virtual card and gesture improvised as the road to investigation. The instrument is an interface, which creates the contact point in the sense of the border between the score card and the sound territory (J.L.Boissier). The instrument as the Chinese scoring board contains the digital memory of the system which produced it. The rule is embodied in the ergonomics. Everything is transformed when model becomes dynamic and utopian. The virtual instruments are at the same moment pure epistemes, systems of knowledges, and instruments of musicology, while the traditional instruments coexist in the same space, divergent and contradictory tensions which contain the aspiration of the various bodies in the sense. The property of the tool, the singularisation "merges" a musical instrument of it in all its depth and its multiple dimensions, to make a body with him, with him, and for that hug him, and opposite: an instrument to measure in future, for the body and the spirit in action. The demanding bassoonists cut their own reeds to adapt itself to their sound body and succeed in reaching their “sound”. A symbiosis is realized between the bodies which are not details exchangeable as in the functioning of the standardized machine. The incalculable, fathomless potential that is the body, the instrument crystallizes it in a tense and acrid report, to remember itself Glenn Gould in the piano on its chair with a hole...

So the instrument is constituted by a program and by ergonomics, a surface which has characteristics of emergence. It is a sedimentation, a superimposition: the bassoon for example contains on its wood, additions, mechanical impasses, scars of failed then resumed technical adventures, as in the improvisation where one can return behind and where errors are turned into qualities, faults are reinstated as accidental alterations which can open new ways. The instrumental space of art coexists and sets up logics, techne and episteme: art and knowledge embodied. Beyond the opposition between techne and logos which is mostly evoked, appears an instrumental conception of the art, which does not presuppose a progress or a model to be achieved. Model can vary, it is dynamic, and its plan is transposable and reproducible in other mediums. It is an opened, characteristic model of multidimensional and virtual spaces.

6. The cartography of a musical instrument

The creative research laboratory “Cartography and networks” of the Ecole Européenne Supérieure de l’Image EESI proposes a multimedia realization which shows the cartography of a music instrument: the bassoon. This research comes within the scope of the art of cartography, a privileged domain of the visual artists from Albrecht Dürer to Jasper Johns, states that a music instrument is a territory that one can explore and represent. A reticular space composed of multiple relations between interpreters, composers, instrument makers and the audience and from which several maps which are visible, readable and explorable by the spectator can be drawn, sensitive and dynamic planes which represent each in their own way the uses, modes and proprieties of the particular interface that is the bassoon. This artistic reflexion and experimentation which finds its inspiration in organology has followed a research we have lead for five years on cartography, data visualization and the notion of network. It was first triggered by the question of landscape and its representation, and then spread to that of the territory and the questioning of networks and applies now to the instrumental space of the bassoon as a network. This proposal has been the topic of a session of experimentations and modelizations with the bassoonist and professor Franck Leblois. This relational space which goes far beyond the instrument as such can be qualified of instrumental. It is composed of the entirety of the bodies which are in an instrumental relationship, the body of the instrument, that of the interpreter and others...

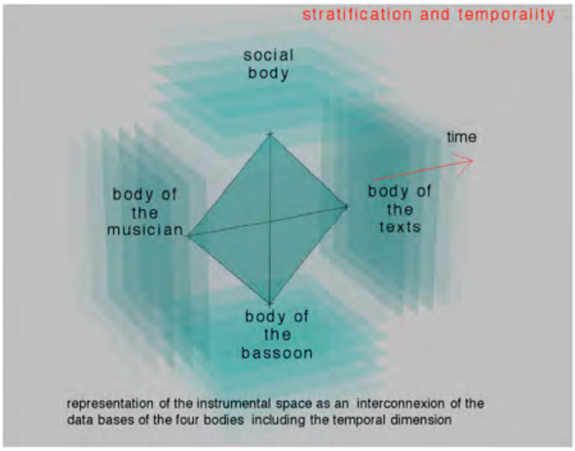


Figure 2. Representation of the instrumental space as an interconnection of the data bases of the four bodies including the temporal dimension.

We shall try a modelization of the intercrossed relationships which compose the global territory of the bassoon which is shown at once from different angles in order to be captured as a whole. This theme of the instrumental space is directly connected to the question of technology within the scope of artistic practices as they are tackled in the Art Schools. This work of research and multidisciplinary creation which deals with the central theme of the tools used in the visual arts at the very day of the digital era will be developed in a first phase during the International Bassoon seminar 2007, at the Gabriel Fauré Conservatory of Angoulême where a very important visual and sonorous database has been gathered dealing with the instrument, the composers, the repertoire, the interpreters, the instrument makers and finally all the protagonists which had met for this unique opportunity : www.eesati.fr/cartographie/ . Here it is not a question of improvising ourselves as cartographers or surveyors, nor to claim to expose a sociological analysis of the instrument, but of limiting ourselves, precisely to our position as artists and so to propose an angle of focus leading to this triple aspect of views, sounds and control which today can characterize a playable position within a space of sound. The other is related with a meta-instrument, a language is forged, within the decompartmentalization of the perceptive system (sight, hearing, touch...). Beyond the known frontiers of audiovisual, the idea of instrumentation by interaction is reconsidered. Our project aims at questioning the concept of hypermedia fitting together of mediums-instrument of an instrument-within the framework of a new emerging category: instrumental interactivity. The "cartographed" territory is studied according to four layers I defined as a reading of the instrumental space: - the musician's body-the corpus of the instrument-the corpus of texts-the social corpus. The presentation of these four readings is carried out from the example of a suite of pieces interpreted by the bassoon, recorded and presented within the terminal as reference notes. So the successive interpretations are so many circuits on the map, paths of exploration on the map, a surface of inscription and spatial representation of the links between the corpuses.

References

[1] Smithson, Robert, “A sedimentation of the mind”: Fourth projects. Art Forum Sept. 1968 reprinted in compilation, with an experience of the exhibition, Paris, 1998.

[2] Couchot, Edmond, La Technologie dans l’Art, Ed Jacqueline Chambon, Paris, 1998.

[3] Boissier, Jean-Louis, Vertus des mondes bornés, Cahiers de l’Ircam N°1, Centre Pompidou, Paris, 1992.

[4] Foucault, Michel, Les Mots et les Choses, Gallimard, NRF, Paris, 1966.

[5] Schaeffner, André, Origine des instruments de musique, EHESS, Paris, 1994.

[6] Vostell, Wolf, Happening & Leben, Hermann Luchterhand Ed.Belfond, Paris, 1976

Human Friction, Material Friction

Jean Marie Georges*, Isabelle Antoine **, Sarah Carvallo*,
Juliette Cayer-Barrioz*, Alain Le Bot*
* *Ecole Centrale de Lyon, LTDS, UMR CNRS 5513, Ecully 69130, France*
** *Pole Supérieur de Design Rhône-Alpes, Lyon 69001, France*
E-mail: georges@ec-lyon.fr

Abstract

The paper confronts the differences of points of view of the concept of friction, which is investigated within the framework of different fields: natural, engineering, human sciences and plastic art. This work, based on several projects involving students, reveals how friction synthesizes all the dissipative processes of energy and information produced in contacts.

1. Aim of the study

The concept of friction appears in distinct fields of human practice and knowledge. Sciences, engineering, human sciences, art deal with it. Interestingly in common sense, friction appears often in a negative way: the unconfessed goal is to suppress friction. At the opposite, the scientific, engineering, philosophical or aesthetic discourses show how it is a constitutive fact. But do they still treat of the same reality, or is their convergence only a coincidence? In other words, is the word "friction" equivocal or univocal? The thesis of this paper states that one cannot understand "friction" phenomenon without dealing with scientific, engineering, philosophical and artistic theories on the subject. By consequence, the word "friction" is univocal but complex. This status will be demonstrated by analyzing – without pretending to sufficiency - some approaches of sciences, engineering, philosophy and artistic creation, which are completing each other.

2. What does friction mean for scientists and engineers?

In classical tribology (the science of solids in contact) [5], friction is the force that opposes the relative motion or the tendency towards such motions of two surfaces in contact. But also, friction synthesizes all the dissipative processes of energy and information produced in the contacts [7], so the friction is related to contacts, dissipations and transfers from one solid to another. The dream for a mechanical, physical or chemical tribologist, which is concerned with friction, is to consider only the contact between two bodies. Fortunately, the reality is more interesting. In the interface between these two bodies, a third body corresponds to the presence of accumulated matter in the contact of the two surfaces. It is the consequence of the surface films, which can flow in the macroscopic contact and control friction. The advent of new analytical tools in surface science, which is the science of the nano quantity of matter, the development of miniature mechanical systems have led to unique approaches at the microscopic and atomic scales. Modern terms, such as microtribology and nanotribology, which imply measurements on different length scales, have been introduced. The methodology begins to provide new insight into the fundamental mechanisms of friction that challenge the classical description of these phenomena [11].

3. What does friction mean for philosophers?

Phenomenology teaches how friction constructs our perception of ourselves, of the others or of the world entities. For instance, the experience of flesh in others or in us radically differs from the way of touching machines or natural things. As pointed out by Husserl or Merleau Ponty – with distinct conclusions –, the feeling of myself arises in the contact of cross-legs which gives the sense of being the same and the other in the mean time. As showed by Sartre, the specific status of the flesh emerges especially in the caress, which reveals the presence of desire. Each time, friction goes beyond utility and materiality to reveal more than the material aspect of things: contacts between human bodies gives substance to a kind of a third body – the flesh, the "ego", the "alter ego" – which expresses itself by specific dissipation effects and desire's transfers.

But we can analyze friction in other dimensions, for example in a social organization. For instance, Clausewitz [1] [2] uses the term friction to describe the effect of reality on ideas and intentions in war. Recent analysis of his work [16] emphasizes the three connected factors: (i) the nonlinearity of combat processes, (ii) the constraints imposed by human physical and cognitive limits, (iii) the informational uncertainties and unforeseeable differences between perceived and actual reality.

4. What does friction mean for the artist?

The rules played by the friction in the work of a Japanese artist Masuo Okabe is taken for demonstration of the many aspects of friction realities and their links [13]. This artist uses a “vestigial" technique: the frottage, whereby the surface pattern of some object is rubbed onto another surface. The origin of this technique can be found deep in the historical past, probably back to the invention of ink and paper. There are painters who draw on the walls of Paris, but few painters, like Okabe make a frottage from those walls [13]. Many cities like Paris, Hiroshima, are used as “gigantic printing plates”. The interesting feature of frottage is that the exact same image never appears twice. It all depends on the different friction processes due to the amount of effort applied, the direction of tracing, the particular habits and features of the person doing the rubbing. As for the automatic handwriting, the frottage is made so rapidly, that the reason and preconceived ideas have no time to apply control. Frottage is thus both the recording of the subject surface, and it is the record of the action of stroking. So, according Okabe « frottage was a means, revealing his own world was the goal »[13].

Friction processes (with contacts, dissipations, transfers) occur not only in the material space, which includes the technique and the formalization of actions. Friction processes occur also in the non-material information space, in the contact of the ideas. The Okabe installation is like a “ library of vestiges or traces”. Each frottage is an element of the system of memory, each link between space and event made as a commemorative stele. The vestigial traces that humans have fostered over the long history of mankind haven been transferred to a nonmaterial stage, and this has freed memory from scriptures and flow through information space where the lost of the past and the preservation of the past is instable.

It has been noted previously that humans use vestigial techniques in the externalization of memory, and this creativity is an approach for revealing the world. Paul Klee said, "Art does not reproduce the visible, it makes visible”. This reference to art is used for the analysis of the ambiguity of friction.

Semantically what are the friction stakes? The meaning of the word friction, which is different for the physicist, the engineer, the philosopher or the artist, induces effects on the use of the word in speech. Is friction a metaphor or a reality at the intersection between the many approaches used in physical and human sciences?

5. Method of analysis

The same project is proposed to two student groups with different academic backgrounds: the first group follows artistic studies (BTS La Martinière**) while the second studies engineering (ECL*). For both groups, the project presents two aspects. A double presentation places side by side, first, the tribology scientific approach, and secondly, the philosophy of contact phenomenology and in particular works taken from modern art. Realizations of paintings and experimental pictures made by the two student groups describe the particular implication of each student.

In this paper, three examples relate the friction approach in natural sciences (such as in physics), to the one in engineering sciences (such as in material sciences), to the one in human sciences (such as in philosophy) and painting arts.

6. The rocking duck

In this example (Figure 1), three points of view appear: the physicist wants to understand natural friction, the material scientist to control it, the artist to express it.

Figure 1: Movie pictures of the duck toy, which comes down the inclined board. The movement is followed from top to bottom in the first and the second columns. The duration between two successive images is about 1s. Its irregular movement is due to slippery displacements of its two feet. Static friction of footboard contacts controls the motion process. This friction is due to the horizontal board angle and to the two materials (board, foot) nature used (here in wood). As an example, if for a given angle, a plastic film is introduced in placed between the duck and the support, the movement is stopped.

The physicist is interested in the trajectory of centre of gravity (X in Figure1) and the influence of all mechanical parameters on movement. The movement of surface asperity is controlled by potential energy of the load in the potential field imposed by the antagonist surface. Dissipations are the origin of friction, as described by Coulomb, Tomlinson, Fraenkel, Kontorova models [12][6][7][11]. This simple toy is a macroscopic illustration of friction theories with all its constituent features of non-linearity multistability and dissipation processes.

The material science point of view is rather different. If the toy is not made of wood but of plastic, the angle at which the duck comes down is different (Figure1). Why it is? The numerous parameters, such as for instance the physicochemical nature of the materials and the surface roughness due to the manufacturing process, play an important role. Thus, the correct angle at which the duck movement starts is found empirically. Frictions experiments are often made without any friction theory. Understanding the phenomenon does not appear necessary in order to control the movement. This example reveals that the techniques used to create an object are not always based on a theory supposed to explain the experience.

The artist integrates the system of the rocking movement into his personal artistic process. How an artist would treat the same phenomenon? We could imagine first that the artist will establish semantic links between the duck and the slippery swing action, will associate other animals able to swing (a chicken, a pigeon) and will develop a significant chain useful for narrative work. We could imagine second that another artist will pay attention to the technical principle of the toy. In the two cases friction is seen as a basic condition.

7. The cell phone

The conflicting viewpoints between the engineer and the artist are clearly visible in the example of the cell phone.

For the engineering designer of such a new technological tool, it is well known that multiple contact keyboards are the weakness in terms of device reliability. The improvement of each contact is a technological challenge. Therefore might it be possible to consider the manufacturing of keyboard less cell phones?

Painting in Figure 2 is organized with multiple prints left by a cell phone. An acrylic red paint layer, deposited beforehand over the phone, is transferred to the white canvas support with a vigorous gesture. The spreading quality of the layer gives visual information about the pressure of the squeezed layer. As an analogy, the paint layer can be seen as an indicator of the contact between the cell phone carried by the hand on the canvas. The great number of indentation spots reveals the importance of these contacts for the painter. A print of lips, at the bottom of painting 2, is like a caressing signature. Transferred layers, which are a physical consequence of friction processes, reveal the human tendency to elaborate technological prosthesis [10].

Figure 2: Painting of Aurélie Tripier. “I wanted to express the traces left by my cell phone, in memory of its existence”.

The cell phone belongs to the field of collective and individual anthropologies. The main engineering interest is for communication in society, and the painting also suggests a crowd. Moreover sociology considers the cell phone as a transitional object submitted to the modes of consumer society, but constitutive of the personal identity, especially for young generations [15] [8]. Thus, if the engineer suppresses the keyboard, the symbolic meaning of the cell phone could be destroyed.

8. The orthodontic apparatus

The design of an orthodontic apparatus, its form, its roughness, and its material surface treatment can be controlled in order to reduce friction against the gum. In best cases, saliva hydrodynamic ally lubricates the apparatus-gum interface. In other cases, painful instabilities are produced. This nonlinear phenomenon is clearly detected by this student as shown in Figure 3.

Figure 3: Painting of Ophélie Bougot. The friction of the orthodontic apparatus against the gum is very badly supported by the painter.

The external object is not well accepted by our body and its artistic expression translates this uneasiness [14]. The pictorial expression (Painting 3) inspired by Francis Bacon and analysed by Deleuze [4] shows deformed features in the middle of the picture, This blurred vision does not describe a temporal vision, but a symbolic haze, corresponding to the juxtaposition of many images, obtained by haptic creation [4] [3]. It suggests how “the friction, which attacks and destroys the painting in fact, destroys me, because I am the painting”.

This example suggests that, on the individual human level, two aspects appear: the friction between one and another, the friction between ego and oneself.

9. Conclusive remarks

Consequently, some objects, strongly determined by friction (the duck toy, the cell phone, and the orthodontic apparatus), have different places and functions. So the transposition from one domain to another (social use, scientific interpretation, technical conception, art representation, philosophical interpretation) reveals the complexity of the object due to the encounter of fields which are commonly separated thus blinding one from the other.

This experiential review suggests that the very concept of friction occupies different places and correspond to different phenomena (technical, scientific, social, ideology, consciousness.), which are treated by natural, engineering, human sciences and plastic arts. So the concept of friction arises at the intersection point of these fields.

This preliminary work encourages us to investigate more in the future. It exemplifies the three points of view upon friction and an educational method for a multidisciplinary approach to technical concepts.

[2] Clausewitz, Carl von, Von der Kriege, trad. fr.: De la Guerre, Naville D. Minuit, Paris, 1973.

[3] Coellier S., “Histoire et Esthétique du Contact dans l’Art Contemporain”, Provence Univ. Pub, Aix en Provence, 2005.

[4] Deleuze G., Logic of Sensation, Francis Bacon, Logique de la Sensation, Seuil, Paris, 2002.

[5] Dowson D. History of Tribology. Engineering Ltd London and Bury St.Edmunds, UK. 1998.

[6] Friedel J., de Gennes P.G., “Friction between incommensurate crystals”, Phil. Mag.87,1,39, 2007.

[7] Georges J.M., Frottement , Usure et Lubrification. CNRS Editions, Eyrolles Eds Paris, 2000.

[8] Gras, A., Joerges, B., Scardigli, V., Sociologie des Techniques de la Vie Quotidienne, L’Harmattan, Paris, 1992.

[9] Heisenberg W. The 1942 Manuscript, Le Manuscrit de 1942, Allia Paris. 2004.

[10] Leroi-Gourhan A., The Gesture and the Language, Le Geste et la Parole, tome II, Paris : Albin Michel, Paris. 1965.

[11] Mazuyer D., Tonck A., Cayer-Barrioz J. “Friction control at the molecular level”, Superlubricity, edited by Erdemir A. and Martin J.M., Elsevier, 2007.

[12] Müser M. H., “Nature of Mechanical Instabilities and Their Effect on Kinetic Friction”, Phys. Rev. Let. 89,22, 2002.

[13] Okabe M., Is There a Future for Our Past, The Dark Face of the light, Minato. C., ed.University of Tokyo Press, 2007.

[14] Sartre J.P., Being and Nothingness, 1943, L’Être et le Néant, tel Gallimard Paris.1976.

[15] Tisseron S. Comment l’Esprit Vient aux Objets, Aubier, Paris. 1999.

[16] Watts B.D. “Clausewitzian Friction and Future War”, Mc Nair paper 68, Institute for National Strategic Studies Washington. USA. 2006.

Acknowledgements

The authors wish to thank R. Mouleyre, D. Mazuyer, J.L. Leonhardt , L. Solana, M. Linard and C. Sherwood for their kind advices.

This research work was supported by the Région Rhône Alpes.

The Nomadic Ear ~ journeys in location aware spatial audio

Nigel Helyer

School of Computer Science and Engineering UNSW Sydney, NSW 2052, Australia E-mail: nigel@audionomad.com

Francesca Veronesi

Key Centre of Design Computing and Cognition, Wilkinson Building, University of Sydney, NSW 2006, Australia E-mail: fver4755@mail.usyd.edu.au

Daniel Woo

School of Computer Science and Engineering UNSW Sydney, NSW 2052, Australia E-mail: danielw@cse.unsw.edu.au

Abstract

This paper introduces the AudioNomad art and science research programme which, in the simplest interpretation, is a location aware virtual and augmented audio reality system. AudioNomad operates with a 'naturalistic' landscape metaphor of our sonic experiences, the software develops a sonic-cartography that co-locates virtual audio with physical features of the environment, allowing a soundscape to be 'performed' by a mobile user or vehicle. This paper highlights some of the challenges encountered during the development of the system and identifies a range of creative and scientific solutions that are here illustrated by two classes of project; location-aware, surround-sound installations mounted on ships and mobile handheld devices for individual pedestrian use.

AudioNomad an Introduction

Audio Nomad comprises a series of cross-disciplinary art and science projects working with the concept of GPS-driven, location-based audio applications. The project outcomes generally take the form of artworks (interactive installations and exhibits) that enable a user or audience to experience a virtual audio world that is intended to be perceived as though it was situated within the real world, so that the user experiences the synthetic sounds as seeming to originate from real objects. Two or three-dimensional audio spatialisation is used to simulate realistic situated sound sources, while non-spatialised sound sources may also be used as location-based content. Conceptually, sound functions to reveal information or create an aesthetic effect. In AudioNomad artworks, the content often entails a combination of oral histories, archival audio, local historical information, field recordings, music related to the general area and specific locations. The outcome is a culturally significant, temporary artifact utilizing this new location-based audio medium ~ an application of global positioning, audio technologies, and software engineering.

Conceptual Origins ~ Tristram Shandy

There is a marvelous passage in "The Life and Opinions of Tristram Shandy" [2] (published between 1759 and 1767) by Laurence Stern, that describes a unique map, made at one to one scale; that is a map made to fit exactly over the physical features it represents! The *AudioNomad* research programme operates a sonic cartography with very similar characteristics, due to the potentially vast scale of the geographic area available to the GPS enabled system and amplified by the fact that the sound composition is performed by the mobile presence of the user traversing real geography.

Conceptual origins ~ Ars Memoriae

Yet another literary source provided the conceptual impetus for the development of a sonic cartography able to seed a physical environment with virtual audio *memories*.

The storage (and retrieval) of audio content within a complex soundscape, virtually associated with real landscape objects, has its precedence in the classical mnemonic system for storing rhetoric. In "The Art of Memory" [3] (pub. 1964) Frances Yates paints a vivid picture of the antique technique that enabled Orators to place memory objects (such as lengthy quotations) within the labyrinthine spaces of classical architecture. By visualizing an architectural interior (real or imaginary) the speaker might place here a red cloak over a sculpture (as a mnemonic trigger) and there, a sword on a table to locate another yet another passage. By memorizing a stroll through this virtual architecture an Orator could retrieve a vast amount of correctly sequenced rhetoric.

The AudioNomad project transmutes such imaginary architectural space into the cartographic space of a digital map (itself a representation of the physical site of the project) and develops a complex sonic landscape by assigning soundfiles, trajectories and other properties, at multiple locations within this virtual domain.

Whereas the classical rhetorician would *re-play* a walk through an imaginary architecture, to sequentially retrieve the elements of a speech, the participant in an AudioNomad project literally walks in a real environment, their position and orientation driving the software soundscape to be delivered to surround enabled headphones. The user experiences the soundscapes as immersive 3D spatialised sound. Audio events appear to be 'located' at specific points in physical space and share similar acoustic properties as the surrounding ambient sounds – forming a seamless nexus between the real and the virtual. Thus, users experience a type of parallel audio world, in which memories of particular sites are invoked alongside contemporary reality.

Alternatively the AudioNomad system can be successfully deployed to massive surround speaker arrays on large mobile platforms as in the case of the ship-mounted works Syren and Syren for Port Jackson.

Syren a shipboard open speaker augmented audio environment

Syren is a shipboard version of augmented audio reality[1] allowing passengers to link the multichannel surround sound experience on board with the visual experience of the surroundings. Geo-spatial information is automatically accessed as the ship navigates the electronic charts associated with the ship's track, via a high-resolution GPS system coupled to a digital compass. This positional information is in turn be used to render a surround-sound, 3D soundscape corresponding to proximate physical features.

Syren was first deployed on the passenger cruise ship *Opera* on the Baltic Sea as part of the International Symposium on Electronic Arts (ISEA) in August 2004. A further developed version called *Syren for Port Jackson* was run over three days in March 2006 in conjunction with the conference New Constellations: Art, Science and Society at the Museum of Contemporary Art in Sydney.

These projects test concepts of virtual soundscape design and compositional processes at a much larger scale than experienced with handheld portable AudioNomad Virtual Audio Reality[2] systems.

Each maritime project is conceived as a specifically designed 'sonic seascape' that positions sound content on, or around, various features of navigational and maritime structures. Special attention being paid to local maritime narratives, histories and specific maritime 'keynote' elements of the local acoustic ecology and seascape. These components are interwoven to form a unifed sonic narrative derived from both ancient epic voyages (as the title Syren implies) as well as contemporary political and cultural life. The soundscapes include ambient and environmental sounds, various music and songs as well as multi-lingual voice narratives drawn from archival and live recordings.

As the ship navigates through and around the harbours the software calls up elements of the soundscape, rendering them in the appropriate direction and distance - simulating a real sound associated with the landscape/seascape.

One of the principal effects of *Syren* is to suggest a series of parallel audio realities that appear to overlay the visual land/seascape and open a possibility to acknowledge a historical and cultural axis pivoting on a geo-spatial point.

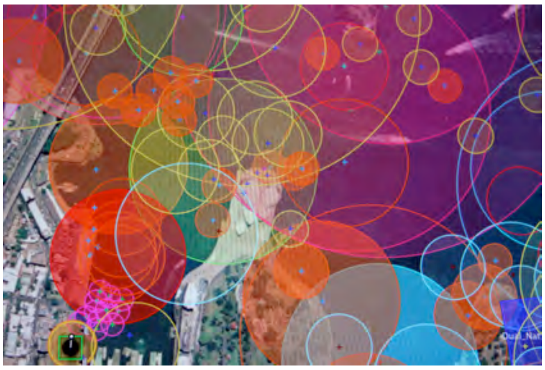


Figure 1. AudioNomad Soundscape composition 2005.

[1] Augmented Audio Reality refers to a system which allows an auditor to experience ambient/local sounds whilst simultaneously overlaying these with additional audio information. The "Audio Nomad" project will operate primarily in a VAR mode whilst the "Syren" project is essentially in the AAR mode (as it is an open speaker system). Both projects will however contain immersive, dynamic and spatialised audio.

[2] Virtual Audio Reality refers to a system that immerses an auditor in a dynamic and spatially active audio environment, which may or may not be linked to a corresponding visual domain (real or virtual). The audio supplied is intended as a total environment and supplants any local or ambient sound.

VAR is not essentially concerned with a functional relationship to events and objects in physical reality, it is best employed in totally VR environments or where there is a desire to diminish or suppress the links between the visual and the aural in the quotidian world (as in the iPod). AAR on the other hand has a vital concern to link synthetic audio events and compositional; strategies with aspects of the physical environment through which the *AudioNomad* is navigating (whilst simultaneously navigating the parallel cartographic/sonographic software).

Technical and Research Objectives

The principal compositional and design challenges have been to work with perceptions of geo-spatial displacement wherein distance and speed transposed to volume and loudness and the large scale of physical architectural and land/seascape features ~ especially when proximate.

Conceptual and Sonic Challenges

The *Syren* project places a strong emphasis on a highly imaginative and creative approach to sound composition and sound design in order to highlight the potential of this emergent field of geo-spatially located virtual audio. Unlike conventional sound-design or musical composition, geo-spatially located audio needs to be highly sensitive to its environmental and architectural context as well as to the fundamentally non-linear manner in which the auditor may interact with the content (in this case via the position, speed and heading of the ship in relation to the marine environment). Pedestrian projects also require deep consideration of the user's behaviour, and unpredictable interaction with the content via position, speed and heading in relation to the architectural/urban environment.

Conceptually and sonically, the principal challenge of the research is to develop a 'compositional' strategy able to deliver a non-linear but coherent 'field' of audio. The system provides the possibility for both (apparently) fixed and mobile audio events, as well as several mechanisms for sequencing sound files in a variety of ways.

Another significant conceptual challenge is the re-conceptualisation of sonic events in the mould of a topology, thus escaping the view of the world (and of sound composition) that is 'object' oriented towards one which is relational and inextricably connected – both through both spatial and temporal axes.

Figure 2. "Syren for Port Jackson" Test rig onboard the launch 'Leander' 2006.

Figure 3."Syren for Port Jackson" Surround speaker array onboard the ferry 'Regal' 2006.

Narrative cartographies as interactions between bodies, memories and places

Mapping footprints, lost geographies in Australian landscapes, a research project currently under development, is a walk through a narrative-scape that temporarily overlays a landscape marked with geographical relics engraved by the Aboriginal clans, custodians of Sydney's northern countries before the European invasion. The site offers the visitor spectacular sceneries of the surrounding bush from a tessellated rock plane punctuated by crops. Part of the experience of knowing the place resides in the search for the engravings scattered all over the terrain. Figures of local animals and mythological heroes reveal themselves to an attentive look, distinguishing them from the natural fissures of the ground.

Due to the dramatic effects of the invasion, the meanings of the engravings, their specific cultural significance, remain to a large extent unknown. Hypotheses identify them as teaching aids, used to pass on local knowledges to the young through rituals and storytelling.

The character of the place, its topography, its lost geographies, its supposed use in traditional times as well as the characteristics of the medium employed to intervene on site inform the strategies of mapping developed by the project.

An interactive sonic map, triggered by the walkers' meanderings through the site, mediates the experience of wayfinding of stories and engravings. This is an immersive experience of seamless continuity between the real and the virtual. As the body is engaged with a simultaneous exploration of the narrative and the physical space, inter-subjective dialogues unfold between the walkers and distant storytellers belonging to other places and other times.

For its reliance on the availability of wireless networks and global positioning systems, the medium enables mobility through place, extending situated places to distant yet co-present spatio-temporal contexts. Hence the site is opened to the encounter with virtual narrators who tell from their own experience stories of places, personal memories and legends from all over Australia. This assembled media-geography renegotiates the relations between the archived and the practiced and challenges the abstractness of maps with the experience of corporeal cartographies.

Rather than distancing themselves from the territory, practitioners enact a journey-based knowledge in subjective ways according to their personal explorations and orientation. As they walk, they randomly interlace narrative passages and create unpredictable narrative cartographies. Thus, the map, resisting the zenith paradigm, returns to what Calvino reminds was its original form: the geographical annotation of a journey [1].

Figure 4. Navigation of Elvina site with handheld device, Ku-rin-gai National Park 2007.

Figure 5. A view of a rock engraving, Ku-rin-gai National Park 2007.

Complexity, Time and Space

The construction and the experience of AudioNomad soundscapes occur within both the time domain and spatial domain. Whilst the experience of audio is by nature temporal and generally continuous, the experience of a spatially constructed soundscape is far less predictable. The content of the work itself frequently references a range of historical and contemporary time periods, adding further epochal complexity.

Although works are, inevitably spatially bounded (albeit often geographically large) and may contain principal vectors (i.e. obvious routes or physical barriers) the compositional structure and strategy does not impose a spatial hierarchy or even propose an explicit spatial structure. Spatio-temporal complexity is amplified by both the pace of walking through the physical landscape (and *ipso facto* through the soundscape) in combination with the temporal duration of individual sound events and their position which may be fixed in absolute space, coupled to a trajectory or positioned relative to the participants position.

References

[1] Calvino, I., Collezione di Sabbia, Garzanti, Milano, 1984

[2] Sterne, L., The Life and opinions of Tristram Shandy, gentleman, Oxford University Press, Oxford, 1983

[3] Yates, F.A., The Art of Memory, Penguin Books, Harmondsworth, 1969

Response: An applied example of Computational Somatics

Todd Ingalls[*] Jodi James[*] Gang Qian[*] Harvey Thornburg[*]
Ellen Campana[*] Stjepan Rajko[*] Dilip Swaminatham[*]
Jessica Mumford[*] Kathleya Afanador[*] Shankar Narayanan[*]
(*)Arts, Media and Engineering Program, Arizona State University, USA
E-mail: todd.ingalls@asu.edu, jodi.james@asu.edu, gang.qian@asu.edu, harvey.thornburg@asu.edu, ellen.campan@asu.edu, stipe@asu.edu,sdilip@asu.edu, jessica.mumford@asu.edu , kathleya.afanador@asu.edu, shankar_instru@rediffmail.com

Abstract

In this paper, we describe our efforts to develop a computational framework based on full-body movement to create embodied interactions in an enactive arts context. Key aspects of our approach are explained including our use of Laban Movement Analysis, in particular in our use of Shape Quality analysis. Another characteristics of our approach include the attempt to find relevant affinities between feedback and full-body movement. We give an example of one implementation of these ideas and some observations on its use.

1. Introduction

A major failing of many interactive systems is the lack of embodiment leading to severe disconnect between one's actions and responses. The term embodiment is used widely across several disciplines, in psychology, neuroscience, philosophy and artificial intelligence. In principle, the term refers to the idea that knowledge manifests itself from the harmonious and interwoven relationship between the mind and body and world. In this sense, an embodied view rejects Cartesian dualism, and instead emphasizes the body's central role in shaping the mind as it moves through the world[16].

In developing systems for expressive interaction, in particular applications that fall under the rubric of arts, we have become sharply aware of the need to consider enactive principles in design and implementation. In our previous work[7][8][9][18], the issue of disconnection and disembodiment from technology was brought to our attention through works that may have been successful from the perspective of an audience but were clearly not so for the performers. We recognized that for the performers in an interactive environment designed for proscenium stage it is extremely difficult to feel connected to the performance environment, which is not surprising considering that these environments focus on audience experience and not the quality of the experience of the performers. To the performers these are not interactive, let alone enactive, environments, but rather, at best, elaborate cueing systems and at worse, detriments to expression. From this problematic issue arose our decision to focus on the quality of the user's of experience as our central goal in current project. In our case we understand the user to have both roles as performer and viewer, or perhaps more important, simply dissolving these distinctions.

Our current goal is to create interactive environments, which are personally meaningful for the user to facilitate engagement, creativity, curiosity and attention. We are doing this through a better understanding the user's perceptions of their experience in the space rather than the outside viewer's perception of what is occurring within the space. We aim to create a positive experience for the user, by designing a space in which users can create movement from an embodied state. Connectivity to the system is imperative for users to be able to create embodied movement. Therefore, as we develop the system we realize that there must be a sense of trust and a level of comfort between the user and the space. Since people have a common sense of connectivity with normal physical space, we have designed our interactive space to be a real, or non-virtual, space enhanced by technological feedback. By utilizing real space for our environment, we give the users a familiar context from which to begin their explorations. In some respect our system is similar to how people interact with their normal environments in that they can manipulate and change their environment based on their interactions. However, in our system we hope users will gain a much greater ability to completely mold and shape their physical surroundings into an environment that is highly reflective of the internal self. We exploit this self-reflective aspect of the system by focusing on the user's internal movement intent as the motivation for the feedback.

This project is a full collaboration between artists, engineers and scientists. At its core this system relies upon enactive knowledge in that the cognitive and sensory experience of being within the designed environments that stems directly from bodily action. Action is analyzed by its component features and by its temporal patterns, thus the system embraces adaptation and change on the part of the user, allowing new behaviors and experiences to emerge over the course of the interaction.

2. Computational Somatics

In an effort to better understand the use of human movement as an expressive medium, we have begun to construct a multimedia interactive art system whose goal is to provide opportunities for fostering creativity, exploring embodied knowledge and information, and developing awareness of our own physicality. Reaching this goal requires us to investigate new ways in which computational analysis can extract meaning-carrying qualitative aspects from full-bodied human movement. Furthermore, we investigate how the results of the analysis can be used to facilitate interactivity between the system in the user.

The first challenge is improving the computational understanding of the meaning conveyed by human movement. Work is being done in this area by others on several fronts[4] and we have decided to focus our efforts with insight gained from Laban Movement Analysis (LMA). Our approach is highly mediated by the rich understanding of feeling and intention (emotional, communicative, proprioceptive) in natural human movement that has been elucidated by creative movement practitioners. The foremost attempt at codifying these aspects is that by Rudolf Laban and[5], called Laban Movement Analysis (LMA). Though LMA originated in dance choreography, applications have ranged far outside that of dance, for instance analyzing the social dynamics of children with Asperger syndrome and learning disabilities[12], helping patients with Parkinson's disease regain function and expression [13], and designing believable, expressive 3D character animations[10]. We adopt a computational approach to LMA, which we term Computational Somatics[15].

LMA is broadly separated into four categories: Body, Effort, Shape, and Space. Body is concerned with the initiation, connection, and sequencing of movement across the body. Does movement originate from the core and spread to the limbs? Is it simultaneous or sequential? Effort encodes expressive dynamics: is the movement tightly bound or free flowing, of strong or light weight? Space concerns how the body relates to its environment: does movement imply a large kinesphere (personal space)? Does the body form lines or planes and in which directions? Shape elicits the form, or forming of the body. A critical component of Shape is Shape Qualities, which pertain to axial motions about the body-centered coordinate system: rising/sinking, advancing/retreating, enclosing/spreading. Shape Qualities seldom exist in isolation; however one quality generally dominates and can be said to characterize the movement. Furthermore, there is no single, prescribed way to exhibit any of the qualities. One may advance, for instance, by walking towards something, by pointing, or simply by craning one's neck forward in a slight and subtle way.

The categories of Laban Movement Analysis, especially Shape and Effort, indicate much about a person's feelings and intentions, both physical (proprioceptive) and emotional. Imagine the following scenario. You encounter a friend in passing and stop to chat for a few minutes. She shares with you the recent events of her life and though her words suggest that her career and personal circumstances are positive and fulfilling, there is something in her body language that communicates otherwise. Usually her posture and movement behavior exudes confidence and gaiety, but this day you do not sense this. Instead, you sense a vulnerability, a smallness, a timidity. What is it in her movement quality that reveals this change of attitude?

3. Affinities

A second challenging problem in constructing a general purpose enactive arts environment is in developing effective strategies for communicating back to the user either the intentions of the system or even more simply what the system believes the person is doing. Since our goal is to create a system that can be used with minimal training yet allows for a great degree of creativity (low floor, high ceilings), we wish to utilize natural affinities that may exist between movement and audio/visual media. While there has been some study of potential conceptual mappings between audio/visual media and full-body movement in performance contexts[5] we are most interested investigating the potential perceptual and conceptual mappings between audio/visual media and one's own full-body movements. In psychology, there has been considerable interest in the role that embodiment plays in shaping how we perceive and conceptualize the world around us[2][3]. Embodiment is likely to play a role in how the user experiences the feedback that is provided by the system. For instance, in our everyday experience of the world, the sounds we perceive get softer as we move away from the source and louder as we move toward it. This experience (and/or biologically determined constraints) might have an effect on how users experience the feedback that the system gives. For example, we might expect that feedback in the form of increasing volume tied to the LMA Shape Quality Advancing would be more effective if the speaker that played the sound was located in the direction that the user was moving toward, rather than in the direction that the user was moving away from. In addition to basic embodiment constraints like this, cultural factors may also play a role. For instance music cognition researchers have suggested that culture provides a set of underlying conceptual metaphors that could underlie understanding of musical sounds[10]. It is possible that the match between the framework provided by these metaphors and the sounds produced by the system might contribute to the effectiveness of the feedback. Culture might also play a role at a perceptual level. As children grow up within a particular culture they have many opportunities to hear the sounds that are preferred within that culture, including music. It is possible that through experience their perceptual system becomes "tuned" to perceive certain musical relations more effectively than others, just as their perceptual systems become "tuned" to the phonology of their native language [17]. A final factor that might play a role in determining which feedback is likely to be effective is the aesthetic preferences of the individual user. In terms of basic operant learning principles, feedback that is aesthetically pleasing to the user is likely to act as a reward, and thus increase the probability of the associated action in the future. On the other hand, feedback that is aesthetically unappealing to the user is likely to act as a punishment, and thus decrease the probability of the associated action in the future. Therefore, understanding the aesthetics of an individual (and users in general) is critical to providing effective feedback. In order to provide effective feedback in practice, it will be necessary to examine how embodied, cultural, and aesthetic factors affect user experience in audio/visual feedback environments. We are interested both in how the user consciously experiences the environment and how the user reacts to it implicitly. Because there is so little existing data, we are beginning to conduct a series of studies investigating the effectiveness of different feedback[1].

3. Case Study Environment : *Repsonse*

In keeping with our focus on developing interactions, we have found it more useful to think of the individual manifestations of our work as *environments* rather than *performances* or *works*. The use of this metaphor helps us maintain the attention to the quality of experience for the user, as he/she inhabits this environment. One such environment we call *Response.*

The theme of *Response* is the manner in which media feedback that is delayed at various timescales can create different sensations ranging from immediate connection, to imitation and finally non-imitative response. The environment also explores how these sensations can be masked by other responses of the system.

This environment uses our Shape Quality (SQ) analysis engine with some basic measure of activity of the user. There are two feedback sub-modules that can act either on the immediate SQ Analysis hypothesis and activity or versions that are delayed through a buffering mechanism. Since there are stochastic elements of the feedback engine the response is not deterministic, but still for the immediate feedback has been found to feel very connected to the users sense of his/her movement.

The first sub-module (we will call *Pulsar)* uses SQ Analysis and activity to alter parameters of a bank of pulsar synthesis [14] generators. Pulsar synthesis is a method based upon the generation of trains of sonic particles. In this case we map the posterior probability of the current SQ hypothesis to various parameters. Advancing/retreating and spreading/enclosing control the range of the fundamental frequency of overlapping pulsarets, with advancing/retreating controlling the lower bound and spreading/enclosing the higher bound of a uniform random frequency generator. Rising/sinking affected the duty cycle of the pulsar generators, causing wide modulations in formant frequencies. Overall activity was also mapped to overall amplitude of the bank of pulsar generators. Because the feedback is moderately complex, the relationship between the feedback and the system analysis becomes more ambiguous. This allows the users to focus more on the quality of the environment and the connection to the feedback rather than the accuracy of the system analysis. In other words, while there is a certain level of awareness that they system is analyzing the movement and that the feedback changes as the user moves through space, because the analysis is not completely transparent the users do not necessarily feel the need to understand the specifics of how it works. Instead, users can focus more on their relationship with the feedback and how the feedback influences their movement, which in turn influences them. The predictability of the feedback for repeated movements lowers with this level of complexity, which can then engage users for longer periods of time. In addition the sound is spatialized so as to be located in surrounding speakers where the activity is being sensed. If delayed versions of the analysis are used this can give a sense of the sound following (or chasing) the user. The intensity of specific color of lights (blue) are mapped directly to the activity level.

The second sub-module (we will refer to as *Glisson*) uses a bank of glisson particle generators[14]. Glissons are short grains of sound that have a frequency trajectory (or *glissando*) over the very short timeframes of the grain. Depending on grain length, the affect can be anywhere from clicks to chirps to large masses of shifting sounds. In this case the glissons are shorter (20-90 ms). The trajectory of the glissons is mapped to the rising/sinking probability of the SQ analysis. Rising movement causes individual glissons to rise in frequency and sinking has the opposite affect. Advancing increases the durations of the glissons while retreating lowers them. In the sub-module white light is mapped to the activity of the user.

We often start of with the *Pulsar* sub-module with immediate feedback to the user. After the user has experimented with this mapping the time delay between analysis in feedback is slowly increased. At some point the users have noticed the delay as being uncanny and disagreeable (we are looking at how to quantify this), but after the delay has reached a certain level the users begin using this delay as a feature to create layers of activity. This often starts with short bursts of movements followed by stillness as the user experiences the delayed affect of their own movement. Past as certain level, particularly if the users moves around in the space, the feedback may appear to be chasing the user and at larger time scales the user may find themselves actually responding to the feedback in an inversion of cause and effect.

As mentioned the *Glisson* module serves a disruptive purpose. If the user becomes accustomed to the delayed feedback the *Glisson* can be activated in immediate feedback mode to possibly draw attention back to the feeling of connection with the system. However, what can result from this is a constantly shifting focus between immediate and delayed response and how this affects the user is something we wish to continue to focus on with this environment.

5. Conclusion

In this paper, we gave an overview of some of the aspects of movement and feedback we are researching to produce enactive environments for full-bodied movement. We also describe one such environment, *Response*, and how this relates to our research focus and questions concerning these types of interactions.

References

[1] Afanador, K, Ingalls, T., Campana, E. “Effects of Tempo on Perceived Speed of Human Movement. Proceedings of International Conference on Computer Music, 518-522. 2007

[2] Ballard, D. H, Hayhoe, M. M, Pook, P. K., et al. “Deictic codes for the embodiment of cognition.”Behavioral and Brain Sciences 20: 723-767. 2007

[3] Barsalou , L. W. (1999). Perceptual Symbol Systems. Behavioral and Brain Sciences 22: 577-660. 1999

[4] Camurri A., Mazzarino B., Ricchetti M., Timmers R., Volpe G. “Multimodal analysis of expressive gesture in music and dance performances”, in A. Camurri, G. Volpe (Eds.), Gesture-based Communication in Human-Computer Interaction, LNAI 2915, Springer Verlag. 2004

[5] Eitan, Z. G., Roni, Y. How music moves: a study of the relation between musical parameters and visuo-kinetic associations. Music Perception 23(3): 221-248. 2006

[6] Hackney, G. Making Connections: Total Body Integration Through Bartenieff Fundamentals, London: Routledge, 2000.

[7] Ingalls, T., Rikakis, T., James, J. , Qian, G. , Olson, L. , Guo, F. , Wong. S. “A Movement Recognition Engine for the Development of Interactive Multimedia Works”, AISB 2004 Convention: Motion, Emotion and Cognition — ConGAS Symposium on Gesture Interfaces for Multimedia Systems, Leeds, United Kingdom (2004).

[8] James, J., Ingalls, T., Qian, G., Olson, L., Whiteley, D., Wong, S., Rikakis, T., “Movement-based Interactive Dance Performance”, Proceedings of SIG ACM Multimedia Santa Barbara, California, U.S.A. (2006).

[9] James, J., Ingalls, T., Qian, G., Olson, L, Rikakis, T., “Lucidity: A Movement-based Interactive Dance Performance.” “LEA-ACM Multimedia” Special Issue, Leonardo Electronic Almanac Vol. 15, No. 3 - 4 (2007). 10 May 2007 <http://leoalmanac.org/journal/vol_15/lea_v15_n05-06/jjames.asp>.

[10] Kalajdziski, S., Trajkovic V., Celakovski, C., Davcev, D., “3D virtual dancer animation,” IASTED International Conference on Modeling and Simulation, pp 20-25. 2002

[11] Lakoff, G., Johnson, M., Philosophy in the flesh: The embodied mind and its challenge to Western thought. New York, Basic Books. 1999

[12] M. Miyahara, M. Tsujii, M. Hori, K. Nakanishi, and H. Kageyama, "Brief report: Motor incoordination in children with Asperger syndrome and learning disabilities", J. Autism and Devel. Disorders, 1997

[13] MotivatingMoves, <http://www.motivatingmoves.com>

[14] Roads, C., Microsound, Cambridge, Mass. : MIT Press,

[15] Swaminathan, D., Thornburg, H., Mumford J., James, J., Rajko, S., Campana, E., Qian, G., Ingalls, T. “A dynamic Bayesian approach to computational Laban shape quality analysis,” Submitted to Advances in Human-Computer Interaction, Hindawi Press, Prepint available. 2007

[16] Varela F., Thompson E., Rosch E., The Embodied Mind: Cognitive Science and Human Experience. Cambridge, MA: MIT Press,1991

[17] Werker, J. F., Tees, R. S. “Cross-language speech perception: Evidence for perceptual reorganization during the first year of life.”. Infant Behavior and Development 7: 49-63. 1984

[18] Whiteley, D., Qian, G., Rikakis, T., James, J., Ingalls, T., Wong, S., Olson, L., “Real-Time Tracking of Multiple People from Unlabelled Markers and Its Application in Interactive Dance”, Proceedings of British Machine Vision Conference, Oxford, United Kingdom (2005).

Rubby; Two Degree-of-Freedom Interface with Haptic and Auditory Feedback

Seung-Chan Kim* Seong-Yong Koo* Seungyon-Seny Lee*† Dong-Soo Kwon*

(*) Human-Robot Interaction Center, KAIST, Korea
(†) Culture Technology, KAIST, Korea*
E-mail: kimsc@robot.kaist.ac.kr, koosy@robot.kaist.ac.kr, seny@kaist.ac.kr, kwonds@kaist.ac.kr

Abstract

This paper describes a mobile interface that makes interaction between human and computer more natural and intuitive. The hemispherically-shaped device with squeezable and elastic housing can be manipulated by prehensile movements of the human hand. The device captures two independent control inputs triggered by thumb or the other four fingers. Also analog value that means how much it is being deformed was also sensed through orthogonally located two flexible sensors on inner surface of device. The device was designed so that the control inputs can be easily captured when applying power grasping. It was found that thumb-driven squeezing is faster than other four digit-driven squeezing motion around 70ms for squeezing process and 80ms for releasing process. The features of the secure grasp and simple control methods allow users to access frequently used functions more intuitively and naturally even in the mobile condition. Vibrotactile feedbacks in the form of hand-transmitted vibration auditory feedbacks were implemented for the input correctness.

1. Introduction

Feedback is an event that gives information about what action has actually been done [10]. In light of the feedback especially in the field of mobile interface, the usage of haptic and auditory feedback is on the increase to give users both reality and awareness of the current situations. These physical feedbacks play an important role in developing human-computer interface. There are essential factors when designing a feedback for hand-held device; *passive haptics* and *human grasping model* which will be described in the following section.

1.1. Passive Haptics

Passive haptics are the technique that incorporates passive physical objects into virtual environments to physically simulate the virtual objects. And it was found that it can enhance the level of virtual presence[4]. From the view point of human-computer interfaces, an interface can be equipped with passive haptic display so that it can convey inherent characteristics of objects which can be hardly implemented by using actuators. One good example of the devices that utilized the concept of passive haptics is *Tango*, haptic device for whole-hand interaction with 3D objects[11]. Users can receive feedbacks by the restitution force from the elastic surface of device.

1.2 Human Grasping Model

Naiper classified hand movement of human hand into two types; prehensile movement and non-prehensile movement[9]. And he categorized prehensile activity into two unique human hand grips; power grip and precision grip. In the case of power grip, the object is held as in a clamp between the flexed fingers and the palm where counter pressure being applied by the thumb. In precision grip condition, however, the object is pinched between the flexor aspects of the fingers and that of the opposing thumb [9].

2. The System

Even though hand gestures are the most informative way to express one's intention, their recognitions are challenging issues because of complicated nature of both static and dynamic hand gestures[8]. Therefore, there must be tradeoff between versatility and recognition rate from the viewpoint of interface design. Mobile devices are also faced with control problems due to the limited spaces. To give more intuitive and natural control ways to mobile devices, we focused on the prehensile movements of the human hand, one of the dynamic hand gestures. Even though power grasping is known as unsuitable grasping method for precision manipulation[3], it is still efficient method not only to exert forces to the object but also to convey one's various intentions especially when the object is an elastic body. The device utilizes the simplified

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

model of squeezing hand motion for the two degree of freedom, or abbreviated as two-DoF, interaction.

2.1. Design Criteria

Many user interfaces are designed to take advantage of hand dexterity because precision grip allow accurate control as discussed in the Section 1.2. *Tango*[11] equipped with high spatial resolution surfaces representative device that utilize dexterous hand motion. The device requires precision grasping method to fully utilize its capability of high spatial resolution. Because precision control inherently requires visual attentions on the device or other feedback channels, the interface utilizing dexterous hand motion is performed better under a static condition. We focused on the mobility while various controls are still available.

2.1.1. Mobility of the Interface

We have implemented the interface that induces users to make power grasp for mobile condition. Stable grasps are essential factors to the mobile devices even though excessive stableness loses controllability. Suitable size and stable grasps make control possible even in the pockets.

2.1.2. Simple Control

Because the mobile devices are different from those of office workstation[7], an even simple task requires more time to complete because of control limitations. To overcome the drawbacks, mobile devices with simple control module such as Click Wheel for iPod or screen-clicking navigation for iRiver - direct click on the display without knobs or buttons - have been developed recently. Direction 1 in Figure 1-(a) shows multi-digit prehensile motion while direction 2 in Figure 1-(b) shows prehensile motion of thumb. The force can be transmitted when the pulp surface presses against the object to be grasped[9]. Considering the fact that there can exist two independent squeezing motions in the power grasp as Figure 1 illustrates, prehensile movements can be utilized for the interfacing method. Because we adopted these motions, there are 2-DoF information input methods in this device. Also this control method do not requires visual attentions as described in the Section 2.1

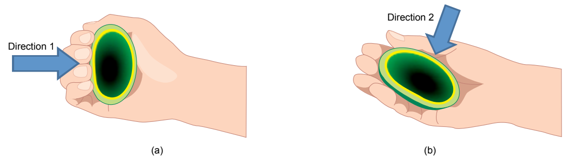


Figure 1. Two different actions during power grasp

2.2. The Device

We have developed the portable interfaces called *Rubby*, named from the material of elastic housing, *rubber*. The device consists of simple flexible sensors inside the elastic housing of approximately 2½ inches(6.35 cm) in diameter and 1 ounce(28.35 grams) in weight. Because the compactness and lightness of the housing, it can be applicable to any kind of interface systems as a control aid when the shape is modified. The sensors orthogonally attached to each other accept two different squeezing motions. These A/D conversion values are transferred to host device through Bluetooth interface. Vibrotactile feedback as a user awareness is conveyed through the vibration motor attached at the center of housing.

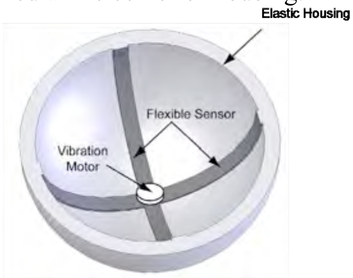
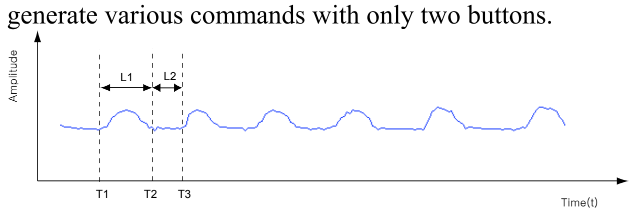


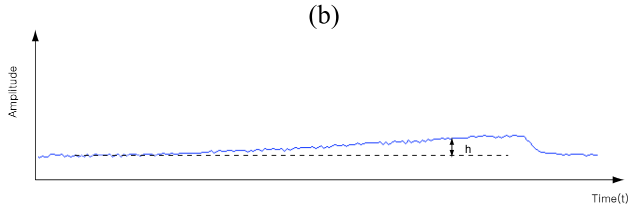
Figure 2. Schematic of the device

2.3 Information Processing

There occurs not only surface deformation but also shape deformation as user squeezes the interface. Because shape deformation is the result of prehensile hand motion, various deformation patterns can be utilized as information input. Figure 3 illustrates one example of prehensile pattern. The time period, marked as L1 and L2, determines action strategies (abbreviated as AS); if L1 is shorter than 500msec, AS is classified as type-1, while L1 is longer than that, AS is classified as type-2. In the case of Figure 3-(b), amplitude of displacement, marked as *h*, is measured so that it can be utilized as analog value. In short, first figure (a), represents digital information gathering method and the second figure (b) represents analog information gathering method. Both of two modes can co-exist. For reference, action strategies stand for various grasping patterns and they can be mapped with various commands as different click patterns(e.g. single click, double click) of the mouse that can generate various commands with only two buttons.



(a) Example of grabbing pattern – AS I, II
(b)



(b) Example of analog value sampling – AS III

Figure 3. Example of the control patterns

Table 1 shows possible actions by the combination of simplified grasping model and action strategy.

Table 1. Combination of Possible Actions

TYPE	AS I	AS II	AS III
Direction I	Digital	Digital	Analog
Direction II	Digital	Digital	Analog

Especially, AS I is similar to mouse click action. It is required to understand the concept of grasping duration and release duration to quantify user pressure-exerting actions. In this research, grasping duration is defined as the time between A, start point of the bending, and B while releasing duration is defined as the time between B, start point of the release, and C, end point of the release action. Figure 4 shows one example of quick squeezing action. For reference, hysteresis effect that implies the difference in the stress-strain relationship during loading and unloading can be seen from the C.

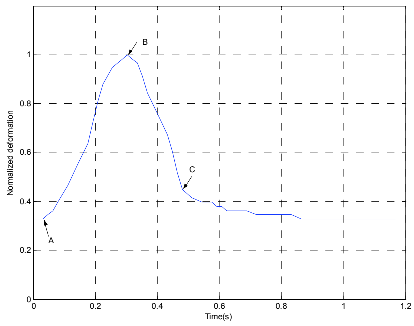


Figure 4. Typical normalized deformation

Figure 5. is the graphical representation of descriptive statics performed by 6 subjects. Every subject was right-handed and guided to squeeze the device as they make mouse click during the operation of the mouse in both two directions illustrated in the Figure 1. All of the subjects are not aware of the object of the experiments and performed two different squeezing motions 10 times respectively. It was found that the duration of thumb-driven squeezing motion was 227 ms with stdev of 41ms for squeezing 233ms with stdev of 44ms for releasing while that of other four digit-driven was 293ms with 34ms of stdev for squeezing 312ms with 52ms of stdev for releasing.

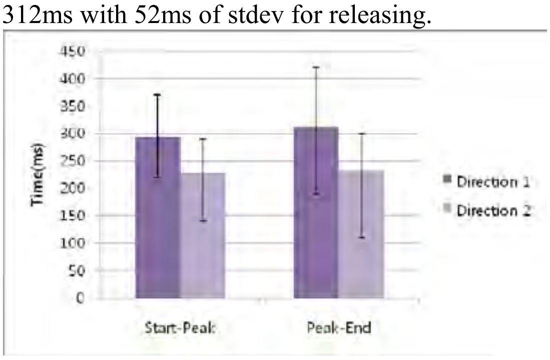


Figure 5. Grasping/releasing duration

The result shows that squeezing motion along the direction 2 is faster than that along direction 1 around 70ms for grasping and 80ms for releasing duration. These results may be utilized when designing squeezable devices.

2.4 Feedback

Because passive haptic feedback does not guarantee whether user motion is accepted as an incoming signal successfully, we had adopted vibrotactile feedback and sound feedback for the user awareness as a complement to feelings by the force of restitution.

2.4.1 Haptic feedback; Hand-transmitted vibration

Because vibrotactile feedback affords additional dimension of perception to visual-information-centered devices[2], it is utilized in many fields such as tactile display, PC-peripherals, mobile devices and so on. In the case of mobile devices, vibrating motor is widely used due to the size and power constraints.

Neurophysiological study reveals that four types of mechanoreceptive units in the glabrous skin area can be categorized as two types; rapidly adapting units (RA and PC units), and slowly adapting units (SA I and SA II units)[6]. Because rapidly adapting units have great sensitivity even though they have poor spatial resolutions, stimulating rapidly adapting units is suitable for generating user awareness efficiently. To innervate the most sensitive Pacinian corpuscle (PC unit) that has the largest receptive fields, pancake pager motor whose maximum frequency is around 150Hz was adopted to generate vibrotactile stimuli. When user action is successfully accepted as an input signal, vibration pulses are conveyed to user's palm. Because desired duration of the vibration had been found between 50 and 200ms[5], 100ms of duration was selected for the vibrotactile feedback.

2.4.2 Sound feedback

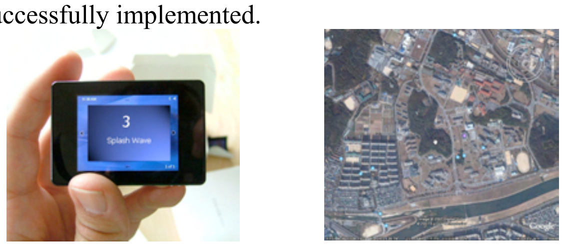
There are three types of sound feedback according to the grasping patterns. When the squeezing motion is treated as a button pressing action, that is digital mode, 1) simple 30ms click sound is provided just after local maximum point depicted as B in the Figure 4. In the case of getting analog value of the deformation, that is analog mode, there are two kinds of sound feedbacks; 2) frequency-modulated sound and 3) a set of beeps at regular interval or gradually decreasing interval. In the case of the former, playing sampling rate is dynamically changes according to the deformation. The sound can be subsonic sound that is barely audible or any kind of application sounds.

$$\text{Playing Sampling Rate(Hz)} = SF_0 \times (1 + dx/dx_{max})$$

where SF_0 means default sampling rate determined by the currently playing sound, dx means current deformation and dx_{max} means predefined maximum deformation. For the case of the latter, short beep like a sound grid of Click Wheel is adopted. The short sound pulse plays as an indicator how fast deformation occurs by the squeezing motion. The interpulse interval is regular for the velocity awareness or gradually decreasing for the spatial awareness.

2.5 Application

There can be two possible application area; a mobile device itself and remote controller for other host devices. First, we have implemented simulator of music player controlled by simple squeezing motion. A user can change playlist, volume, and playing sampling frequency even in mobility condition as discussed in Section 2.1. The small size and simpleness of the manipulation makes controls possible even in the pocket because it does not require any visual attention on the device during manipulation. Second, simple grasping movements were mapped to arrow keys, buttons on a computer keyboard, as Table 2 illustrates. Using the 2D control functions, navigating Google Earth shown in Figure 5-(b) was successfully implemented.



(a) MP3 Player (b) Google Earth

Figure 5. Example of the implemented 2-D navigation

Table 2. Combination of Possible Actions

	AS I	AS II	AS I	AS II
Direction 1	Page Up	Home	↓	↑
Direction 2	Page Down	End	→	←
Application	Presentation controller		Media Controller(e.g. for changing channel, volume)	

3. Conclusion

Considering the number of operations to access repetitive tasks is important factors in the mobile devices[1], simple squeezing motion can give alternative way to control the mobile devices naturally and fast. Because not only secure grasp but also simple control method was achieved, it is expect that the mobile devices can be easier to manipulate even though rigorous psychophysical tests are remained. Also the advantages of multimodal feedback for both fully-sighted and visual impairments[12] can enhance the performance of mobile interface. Furthermore, it is also expected that the suggested interfacing method can be extended as a new emotional interaction method because intuitive and dexterous motion of human hand can be accepted as an information channel.

References

[1] J. Gong, P. Tarasewich, Guidelines for Handheld Mobile Device Interface Design, Proceedings of DSI 2004 Annual Meeting (2004).

[2] R.G. Hughes, A.R. Forrest, Perceptualisation using a tactile mouse, Visualization'96. Proceedings. (1996) 181-188.

[3] K. Imazeki, T. Maeno, Hierarchical control method for manipulating/grasping tasks using multi-fingered robot hand, Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on 4 (2003).

[4] B.E. Insko, Passive Haptics Significantly Enhances Virtual Environments, University of North Carolina, 2001.

[5] T. Kaaresoja, J. Linjama, Perception of Short Tactile Pulses Generated by a Vibration Motor in a Mobile Phone, Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems-Volume 00 (2005) 471-472.

[6] M. Knibestol, A.B. Vallbo, Single unit analysis of mechanoreceptor activity from the human glabrous skin, Acta Physiol Scand 80 (2) (1970) 178-195.

[7] J.A. Landay, T.R. Kaufmann, User interface issues in mobile computing, Workstation Operating Systems, 1993. Proceedings., Fourth Workshop on (1993) 40-47.

[8] A. Malima, E. g, M. tin, A FAST ALGORITHM FOR VISION-BASED HAND GESTURE RECOGNITION FOR ROBOT CONTROL, Signal Processing and Communications Applications, 2006 IEEE 14th (2006) 1-4.

[9] J.R. Napier, THE PREHENSILE MOVEMENTS OF THE HUMAN HAND, Journal of Bone & Joint Surgery, British Volume 38 (4) (1956) 902-913.

[10] D.A. Norman, The design of everyday things, Basic Books, 2002.

[11] D.K. Pai, E.W. VanDerLoo, S. Sadhukhan, The Tango: A Tangible Tangoreceptive Whole-Hand Human Interface, Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems-Volume 00 (2005) 141-147.

[12] H.S. Vitense, J.A. Jacko, V.K. Emery, Multimodal feedback: establishing a performance baseline for improved access by individuals with visual impairments, Proceedings of the fifth international ACM conference on Assistive technologies (2002) 49-56.

The interactive flower:
The optical fiber flower controlling sound and music

Yong ho Kim* Seungyon-seny Lee‡
Graduate School of Culture Technology
Korea Advanced Institute of Science and Technology,
E-mail: breadkim@kaist.ac.kr ,*
seny@kaist.ac.kr‡

Abstract

In this paper, we challenge to provide more active listening environment which enables the listener to control various audio effect and music selection. We made the novel hardware interface which we named '*interactive flower'* using optical fiber lamp as an expression part and water cup as a controlling part. *The interactive flower* used tilt sensor, geomagnetic sensor for detecting tilting and rotating gestures of the cup to control 2-D panning for midi file, the file selection for wav file and the audio effect such as flanging, volume, reverberation.

1. Introduction

1.1 Motivation

There are a lot of reasons that people listen to the music with personal audio device. Some people listen for refreshing their mind, some people for learning how to play their instrument, some people for dancing, and even some people for sleeping. If we consider these various desires to listen to the music, we can easily find out there will be many possible ways of listening music, too. For example, people who want to dance with hip-hop music might want to hear the beat more than the melody, or people who want to learn playing guitar might want to hear only the difficult guitar solo part more than any other part. But, these days, we have only one way to listening music which is just playing recorded musical media with their audio device and passively listen it. [1] We want to solve this problem and give the listener more possibilities of active listening. However, the today's hardware interface of personal audio device, such as a button or a wheel or a slider, looks no longer acceptable for controlling varieties of audio effects or music spatialization, and for more possibilities in an aspect of entertaining interaction.

So, we suggest that the novel interface to control their personal sound environment while they listening music. Besides, which looks more beautiful and has entertaining interaction.

1.2 Related works

There are some related works designing the interface for controlling sound environment. In 1998, F.Pachet and O.Delerue developed the system controlling the location of virtual sound source in real time while they listen to the music with midi file. [1] In this work, they made the 2-D and 3-D software interface for controlling each instrument's location and compare the usability. Their basic concept that listeners can be given some degree of control on the music they listen to is same as our work, *the interactive flower*. But the point that they focused only software environment and we focused physical hardware interface more than software is different.

Another example is M.Hamanaka, and L.Seunghee's paper named 'sound scope headphone'. [2] In their paper, they tried controlling audio mixer by natural movement using of a headphone with 3 different sensors. They said the existing audio mixer has too complicated to operatie for users in general, so they made the headphone which is attached digital compass, tilt sensor, and infrared sensor. After seeing these works, we consider about the intuitive gesture we will use. Finally, for our purpose of making more active listening environment, we chose the natural motion of pouring water from the cup and rotating the cup, instead of other gesture.

2. Concept

2.1 Form concept

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

The first inspiration while we decided the form of *the interactive flower* was from the real flower reaction and applying to an automation flower. Like the most of the other plants, we have the image of flower that is ‘silent’ and ‘standing’. Flowers neither dramatically react via human’s gesture, nor generate sound or light by themselves. Actually, they are interacting with their surrounding, but the reaction is too slow that we can see. For this reason, there’re many imagination and motivation that we wanted to make the flower reaction. We wanted to break these kinds of stereotype that most plants have. Why it can’t be ‘moved’ or ‘controllable’?

So, we designed *the interactive flower* reacting opposite way to the real flower. *The interactive flower* react dramatically to the human movement, and it generates light itself and controls the audio effect. The pouring gesture which can be detected by tilt sensor corresponds to the motion of pouring water into the flower, and the smooth growing motion of real plant corresponds to upward and downward motion of *the interactive flower*.

Actually, there were a good example of robot flower reacting to the human’s motion which called Cyber flora by MIT media lab. [3] These robot flowers detect human body temperature using heat sensor, and close its petals when people put his or her hand near the flowers. Also it opens and traces warm human body when people go away. We added some element of the light to the concept of ‘moving flower’ using optical fiber with RGB LED.

The optical fiber which represents the petal of *the interactive flower* has been main material for effective little energy loss optical communication. But, lately, people start to focus its value of light display, and start to use variously as a lighting material for lamp exhibition, fine art, or the light source. [4] The optical fiber can be separated into end light type and side light type. *The interactive flower* used end light type plastic optical fiber to transfer vivid color of LED light source.

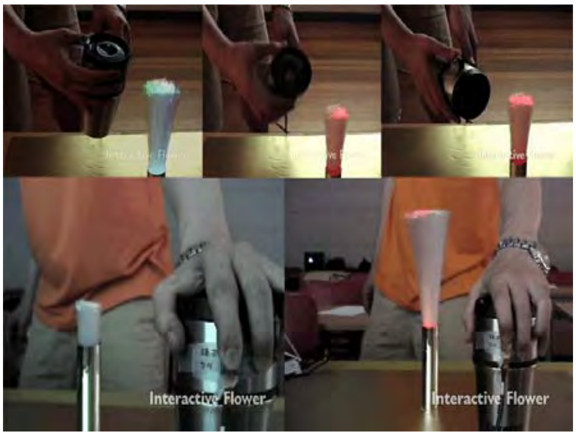


Fig1. The appearance of *the interactive flower*

2.2 Function Concept

Every flower’s motion and audio control of *the interactive flower* made from the communication along with the Micom module, the Sensor module, and the Max/msp software [5]. If the listener tilts or rotates the cup, the sensor module below the cup which contatins tilt sensor and geomagnetic sensor detects the tilting or rotating motion, and it sends the signal to the micom module. The micom module provide the electric voltage to LED light source and servo motor which makes the increasing and decreasing vertical motion of *the interactive flower* at the same time. The signal from the sensor module transferred to the Max/msp patch, too. And Max/msp analyzes the sensor signal, matches with 2-D panning and other extra audio effect. Thus, the tilting and rotating gesture is connected with the audio effect control.

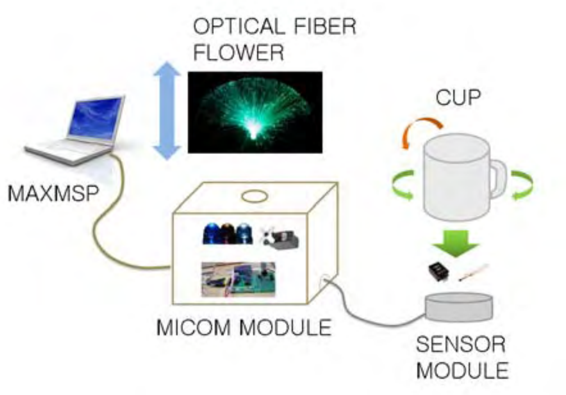


Fig 2. The Basic Concept of *the interactive flower*

2.3 Scenario

Our imagination became a short scenario, and the main story is very simple like below.

When we pour the sound to the flower, it will grow up having light and making sound. And we rotate the cup, the sound gives us the spatialization. At first, a user selects music file on the Max/msp patch. If they select the midi file, they can select each instrument sound track, and they want to control and hear the instrument sound track closely or from far away. For example, if they want to hear the violin sound more than any other instrument, they have to choose the violin sound track and tilt the cup to hear the track closer and rotate the cup to rotate the sound. And whenever they pour the sound to the flower, that is, whenever they tilt the cup downward, the flower grows upward, so users can be given visual feedback of their motion, and easily get accustomed to the connection between pouring sound and growing flower.

If they choose the wav file, they can select previous or next music file by rotating the cup clockwise or counter clockwise direction. And at this time, if they pour the sound to the flower, the audio effect such as volume/reverb/flanging will increase with the flower’s growing motion. The type of the audio effect they want to control can be chosen in the Max/msp patch before tilting the cup, either.

3. Audio Control

3.1 Gesture mapping for midi and wav file

Now, we will describe the concrete information about the audio control that we introduced in *the interactive flower*. First, we consider two kinds of music file format, which are midi and wav file. Most of common music file format is prerecorded without the information about the angle (azimuth) and the distance between the listener and the sound source, so it can be said that there will be little meaning controlling panning effect with these kinds of file like mp3, wav, FLAC(free losesless audio codec) etc. So, we had to use music data for 2-D panning the midi file or RWC Music data base[8] (the archive of the raw music data before mix-down). We chose midi file this time, thus we mapped sensor signal to the 2-D panning for the midi file, and for the wav file, we introduced another function for selecting music and controlling audio effect.

Table 1. The gesture mapping of *the interactive flower*

File / Gesture	Midi file	Wav file
Tilting	Panning Distance	Audio effect (Volume/Flange/Reverberation)
Rotating	Panning direction	Music selection
Tilting	Growing and falling motion of the flower	

3.2 2-D panning

In natural environment, we get the information about pitch, loudness, timbre, temporal organization from the sound we hear. Beside that we can often decide to a fair degree of certainty from what direction a sound comes, how far away the source of the sound is, whether it is moving and how. [6] This kind of information related to the distance and the direction of the sound source have been developed by people who study the musical perception and psychoacoustics for a long time. And people who study computer science integrated the inter-aural model of hearing sound into the computer music. Nowadays, we can hear the sound which the virtual sound source is moving through the virtual trajectory in the virtual sound space. *The interactive flower* is the controller for moving the location of virtual sound source in the virtual space.

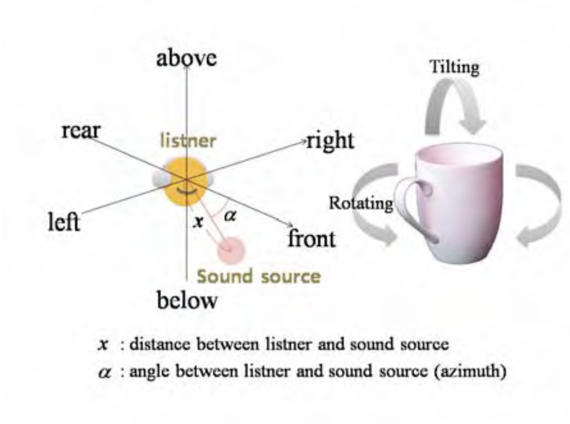


Fig 3. The sound source in the virtual sound space

As we mentioned, we divided file format into two types, and 2-D panning is for the midi file. We used each midi track as a sound source which represents each instrument's musical melody of instrumnet. So, listeners can change the virtual location of each instrument by rotating and tilting motion. If they tilt the cup, the specific midi track or the specific instrument they choose will get closer(x will be decreased), and if they rotate the cup clockwise direction, the virtual sound source will move draw the circle around the listener(alpha will be increased). This interaction can give some degree of freedom to the listener. They select a preferable instrument and the closer located-sound-source will be more clearly distinguished than any other sounds.

3.3 Music selection

For the wav file, we select 12 samples and locate them in the circle which has the same distance around the listener in the virtual music space.(Fig 4) then we mapped the geomagnetic sensor's signal to the orientation of the music file in the virtual music space.

The virtual location of the file which is playing at present is always the front position of the listener, and if the geomagnetic sensor's signal exceed the (+) direction threshold, the next file will be located front position, and the file which previously played will rotate -30 degree. It will iterate the same sequence until the cup came back to the initial position. And because the file would be playing, while the angle changes simultaneously. So, as time flows, the next file will be automatically located the front position of the listener without rotating the cup. In this case, the circle of music files surrounding the listener can be regarded as a time line of the music file.

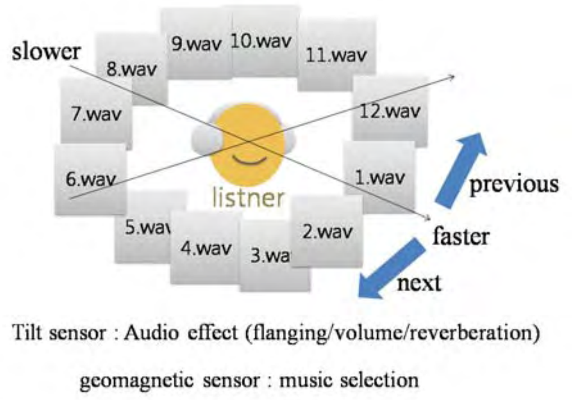


Fig 4. The circle of music file around the listener

3.4 Audio effect

Because we map tilt sensor's signal of the music file to the audio effect such as flanging, volume, reverberation in the case of wav file, listeners can give the effect to the musical part that they want to add.

Flanging is a time-based audio effect that occurs when two identical signals are mixed together, but with one signal time-delayed by a small and gradually changing amount, usually smaller than 20 ms (milliseconds). This produces a swept 'comb filter' effect: peaks and notches are produced in the resultant frequency spectrum, related to each other in a linear harmonic series. Varying the time delay causes these to sweep up and down of the frequency spectrum.

And Reverberation is the persistence of sound in a particular space after the original sound is removed. When sound is produced in a space, a large number of echoes build up and then slowly decay as the sound is absorbed by the walls and air, creating reverberation.

By simply tilting the cup downward, the degree of audio effect will grow. The tilt sensor's signal also connected to the servo motor's motion. So, the flower will grow more and more when we tilt the cup for the great degree. So, we can say that flower's height is expressing the status of the audio effect on the music file. Actually, this expression part somehow looks too simple to show every aspect of the sound environment of the music. But, it does an important role as a visual representation, because with this physical visual representation, the listener doesn't need to check the computer screen, they just can see the status with the physical object. And because optical fiber has the possibility of color representation, it can be a good object for the physical visual representation of the sound environment.

4. Conclusion and future works

In this paper, we proposed the hardware interface for controlling personal listening environment and introduced the concept and function of each part of the prototype interface, named '*the interactive flower*'. With the *interactive flower*, the listener can experience locating the each instrument, or the music file in the virtual sound space and controlling audio effect by easy tilting, rotating gesture. In future, listening environments can be greatly enhanced by integrating relevant models of musical perception into musical listening device, and the role of new audio interface became more important. Especially, in the virtual world, to keep pace with the high speed development of the visual 3D representations, more intuitive and comfortable 3D sound interface also has to be provided. Therefore our next step will be developing 3D sound spatialization control interface which enables individual more degree of freedom to 'active listening'.

Reference

[1] François Pachet & Olivier Delerue(Sony CSL SONY CSL Paris), "A Mixed 2D/3D Interface for Music Spatialization", First International Conference on Virtual Worlds, Springer Verlag Lecture Notes in Computer Science 1434, pp. 298-307, 1998.

[2] Masatoshi Hamanaka (PRESTO, Japan Science and Technology Agency) , Seunghee Lee(University of Tsukuba Graduate School of Comprehensive Human Science), "Sound Scope Headphones: Controlling an Audio Mixer through Natural Movement"

[3] Lieberman, J. CyberFlora exhibition at the Cooper-Hewiit Museum, New York, 2003. Informationavailable at http://robotic.media.mit.edu/projects/cyberflora.html

[4] 'Young-won technology' optical fiber introduction website http://www.light.co.kr/index.htm

[5] http://www.cycling74.com/products/maxmsp

[6] Rossing Moore & Wheeler, The science of sound, third edition 2002, pp.685-691

[7] http://staff.aist.go.jp/m.goto/RWC-MDB/rwc-mdb-p.html

At and Across: Physical and Virtual Action-based Music

Juraj Kojs
University of Virginia, Charlottesville, VA, USA
E-mail: kojs@virginia.edu

Abstract

This paper documents an instance of action-based music composition, which means that physical and virtual actions are the primary carriers of musical meaning. "At and Across" is a piece for physical and virtual Slovak sheep bells. Cyber bells designed by physical modeling synthesis effectively augment the properties and actions of the physical instruments. Furthermore, the models created in GENESIS software [2] enable actions impossible in the physical world. The paper describes instrumental, structural, and timbral areas of the composition.

1. Introduction

Composing with the sounds of mechanical actions in electro-acoustic music has a longstanding tradition. The pioneers of *music concrete* such as P. Schaeffer engaged sounds of everyday activities in their electronic music tape compositions [7]. The composers used technology to manipulate everyday sounds for the purpose of constructing abstract musical structures. Despite a variety of manipulations, the original production mechanisms that generated the sounds remain perceived as they frame the compositions.

Sound production involves sources in a cause and effect relationship. The British ecological psychologist W. Gaver points out that we indeed can listen to the world and music in the 'everyday listening' mode [3], [4]. As opposed to 'musical listening' which favors psychoacoustic perception of sound as it reaches the ear, 'everyday listening' highlights the identification of sources that produced the sound.

In acoustic music making, a performer and instrument present the sources which enable sound production. In fact, the sounds only materialize when the two agents join in action. Bowing, plucking, blowing, scraping, and hitting are some examples of action mechanisms a performer can apply to an instrument.

Action-based music emphasizes that actions are heard as the principal means for musical expression. For example, the compositional focus on investigation of bowing mechanisms can lead to a development of a coherent vocabulary, from which a musical meaning emerges. As with any language, action-based music occurs when the vocabulary is structured into musical phrases and sentences, and is, thus, developed and sustained.

Musical vocabularies of hitting, rubbing, scraping, bowing, blowing, and plucking have their own morphological characteristics, however, they can all be identified and categorized as the lexica of one language: the language of action-based music. Naturally, connecting this multiplicity of vocabularies in a single composition reinforces the perception of action in music.

In the digital domain, physical modeling synthesis facilitates simulation of action and sound production mechanisms. This synthesis approach enables us to imitate and extend the existing action mechanisms beyond the limitations of the physical world. Furthermore, we can simulate instruments and actions which do not exist in the physical reality. Physical modeling is therefore the most effective technique for creating action-based music in virtual space.

At and Across is an electro-acoustic composition which combines actions of physical and cyber resonating structures. In particular, the composition investigates the expressive potential of physical and virtual bells. The following sections document the compositional process including its formal, musical, and technical details.

2. *At and Across*: Motivation

At and Across is orchestrated for a set of physical and virtual Slovak sheep bells. The composition situates the ancient folk instrument in the domain of cutting edge technology. The project developed from the idea of digital excavation and preservation of the Slovak sheep bells.

Sheep bells arrived in Slovakia during the shepherd colonization in the 13th century by the Wallachians of Romania. In the northern region Liptov, where I grew up, sheep bell production has flourished since the 16th century [9]. Following 20th century industrialization and disappearance of small farmers, sheep breeding and shepherding culture and its sounds have been vanishing from the Slovak countryside.

From my early childhood I remember the sounds of the clinging bells at my grand father's farm. All the bells of his sheep were tuned to one common tone. This way the shepherds, who pastured and guarded the

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

animals in the spring and summer, could identify the sheep in case they mingled with other herds which were tuned to other tones. Consequently, it was important for the farmers to get a set of well-tuned bells. A resonant bell was often more valuable than the sheep itself.

Slovak sheep bells were first used in *Zvonenie (The Ringing*, 2007), a composition for cyber bells on 4-channel tape, which was the product of my research investigations with GENESIS at ICA—ACROE center in December 2006. GENESIS is a composer-oriented interface designed for building physical models within the CORDIS–ANIMA environment [1]. The environment is based on the mass-spring-damper approach and enables a creation of modular, freely re-combinable, virtual structures.

The process of composing *Zvonenie* began with study of a physical sheep bell's acoustics. The spectrum of the bell became the primary source for the horizontal and vertical structures of the composition. The digital simulation of the bell in GENESIS and extending its properties beyond the limitations of the physical world followed. Interactions between the virtual exciters such as beaters, bows, and plectra for hitting, bowing and plucking generated complex rhythmic sequences. These sequences were exported as audio files to a digital studio, where the formal compositional work took place. Application of dynamic spatialization and reverberation in MAX/MSP completed the work [6], [8].

At and Across expands *Zvonenie* while engaging a physical performer and a set of physical sheep bells with the orchestra of cyber bells. Furthermore, this composition investigates additional performance modes such as blowing, combines the cyber bells into a series of networks, and renders the melodic and harmonic trajectories more dynamic.

3. The Physical Bells

A set of sheep bells used in *At and Across* was acquired from the bell-making master Julius Mikulas from Ilanovo, near Liptovsky Mikulas. Eight bells are tuned to G mixolydian scale, which is the predominant melodic mode in Slovak folk music. The dominant resonant frequency of the bottom scale tone is G4 (784Hz), while the top note shows G5 (1568Hz) as its predominant resonance. The G4 bell bell is shown in figure 1.

The bells were made by hand from a single piece of brass alloy approximately 0.1 cm thick. The brass was folded and enclosed on its sides to form the bell's cavity. The beater and handle were later welded onto the bell's body. Body size suggests initial pitch consideration: larger bells produce lower fundamentals and vice versa. Hammering around the bell's rim is a technique used to refine and finalize the tuning.

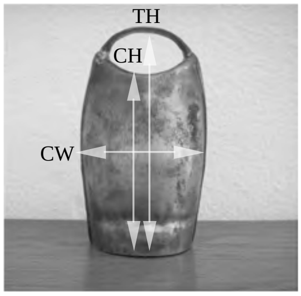


Figure 1. Slovak sheep bell with the following parameters: Total Height (TH)= 13cm, Central Height (CH)= 10.3cm, and Central Width (CW)= 10.5cm.

The bell's weight is about 200g with additional beater weight of around 17grams. The opening has a shape of an oval approximately 6cm wide and 6.3cm tall.

A strong fundamental frequency and a small number of quickly decaying partial characterize the spectral behavior of a brass bell. A single hit on the bell's rim—the most resonant excitation place—produces a piercing attack sound which attenuates within approximately two seconds.

4. The Cyber Bells by Physical Modeling

The physical models were designed with mass-spring-damper algorithm in GENESIS. The model consists of 11 masses of different weights which decrease symmetrically as they move away from the central point. The masses are connected with links of variable, and generally low, damping and viscosity values.

The model was replicated 11 times and tuned to 11 different fundamentals. These fundamentals are related to each other according to the ratios proposed by J.C. Risset in his additive synthesis of a bell [5]. An instance of the model hit by a single beater mechanism is displayed in the figure 2. The beater mechanism consists of a single mass, generally of small weight, attached on one side to the fixed point and on the other side to the bell. While the connection between the beating mass and fixed point is rather rigid, the link between the beater and bell is less damped in order to enable rebounding of the beater.

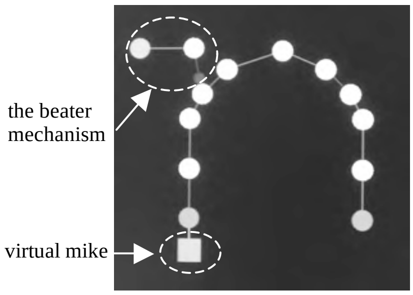


Figure 2. A cyber bell by physical modeling synthesis hit by a single beater in the GENESIS bench.

The model proportions remained constant throughout the composition. The cyber bells were either connected into networks or excited by unusual actions. The network links were designed to disturb the natural behavior of the resonating instruments, while avoiding the collapse of their internal structures. An example of such network is shown in figure 3.

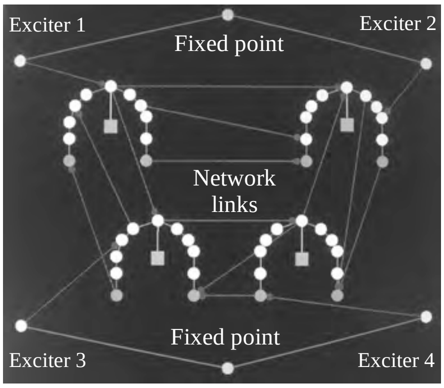


Figure 3. A network of four cyber bells hit by four exciters at different places.

The exciters were modeled as a combination of masses attached to fixed points and linear and non-linear links. Their interactions with the bell structures are described in the following section.

5. Virtual Interactions

Hitting, plucking, bowing, and blowing mechanisms were used to excite the bells. As an extension to what can be done in the physical world, the bells were exposed to unrealistic excitation situations. Multiple exciters of varying kinds acting on the same place of a single bell and single beaters and bows performing on multiple instruments exemplify such unusual excitations.

Figure 4 shows a single exciter bowing the bell at 11 different places. Each bowing-link varied slightly in its parametrical design. Moreover, the positions of the virtual microphones were eventually altered while all other parameters were retained. Figure 5 shows four virtual microphones scanning the bell at four different places, thus suggesting a 4-channel audio “ear”.

In summary, GENESIS enabled me to reach beyond the limitations of the physical world while (a) varying the properties of the exciter and interaction mechanisms, (b) creating impossible performance actions, which resulted in complex rhythms at extremely rapid tempos (c) and simulating unrealistic performance modes such as plucking the bell. These actions resulted in a creation of unique timbres. The sonic identity of the bell and the individual sound production characteristics were nevertheless preserved.

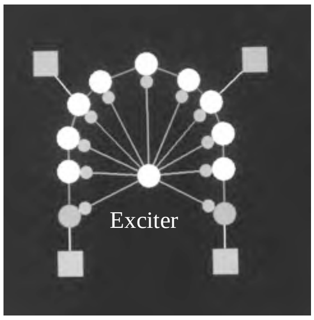


Figure 4. Single force bowing the bell at 11 different spots.

6. Form

Composed interactions were exported as stereo and quadraphonic audio files to a digital studio, where they were assembled into a 4-channel soundtrack. The formal arrangement of both computer and acoustic parts follows the trajectory of hitting, bowing, blowing, and plucking sequences.

The process of action engagement proceeded as follows. First, the sound production mechanisms on the physical bells were investigated and catalogued. The actions included (a) shaking, (b) hitting with metal, wooden, and plastic beaters (mallet and stick parts), (c) bowing with the violin bow, and (d) blowing inside the instrument. All these actions were performed while handholding the instrument in the vertical and horizontal positions. The bells were also hung on a horizontal rod, which enabled controlled performance of multiple instruments. The actions were performed with and without partial or full muting.

Second, the individual actions became the primary framework for the component sections. The piece was composed as a sequence of actions with the following arrangement: *Intro, A, B, A’, x’, C, B’, C’, x’’, B’’, D, x’’’*, in which A= hitting, B= bowing, C= shaking, and D= combined modes. The performer holds all the bells in his/her hands and shakes them throughout the introduction. Lowercase *x* parts signify shorter solo electronics sections. The complete formal trajectory of *At and Across* is shown in figure 5.

7. Interactions

Interactions between the physical and virtual bells were pre-composed and based on the melodic and harmonic plans. As the piece is dedicated to the memory of my grandfather, the composition’s melodics is derived from his favorite Slovak folk song “Sadla muska na konarik” (“A little fly landed on a little twig”). Re-composed segments of the song appear in both physical and cyber bell parts. Each section of the composition is built around a fragment from the song. Depending on the performance mode and

Time	0—2'	2—3'	3'—4'	4'—5'	5'—5'30"	5'3"—6'30"	6'30"—7'30"	7'30"—8'30"	8'30"—9'10"	9'10"—10'10"	10'10"—11'10"	11'10"—12'
Part	intro	A	B	A'	x'	C	B'	C'	x''	B''	D	x'''
Genesis		Hit single bell	Bow	Hit multiple bells	Poly-phony1	Hit multiple plates	Bow, Blow	Hit single plate	Poly-phony2	Blow, Pluck wood	Hit multiple plates and bells	Polyhonny3
Performer	Shake	Hand held (HH) + outside beaters	HH + Bow top	HH + inside beaters		HH + shake	HH + bow + bottom	HH + muted shake		HH + bow all and blow	HH + shake + suspended polyphony with beaters	→

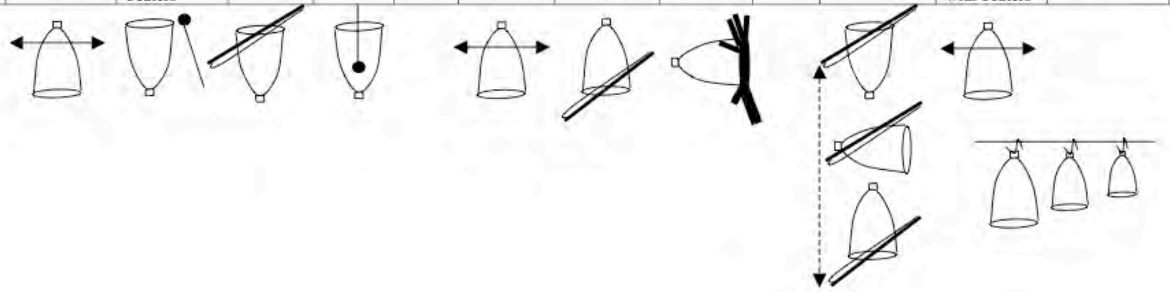


Figure 5. The complete formal trajectory of *At and Across.*

rhythmic complexity, these fragments may not be necessarily recognizable.

Harmonic and dynamic structures mirror and accentuate the spectral relationships between the partials of J.C. Risset's additive synthesis bells. The quartic function—whose inverse version can simulate the decaying behavior of a bell sound—was used as a formal agent for constituting the growing appearance of the cyber bells in the course of the piece.

While the background layer of the computer part flows with its own dynamics, the foreground is strictly structured to complement the physical performance. The final electronics part and acoustic signal joined in MAX/MSP, where they were cross-synthesized and dynamically spatialized.

8. Conclusion

At and Across is an instance of action-based music composition. Such compositions highlight sound production as the primary tool for musical expression. Physical models by mass-spring-damper approach enable a composer to designs modules of sound production mechanisms and interactions which reach beyond the limitations of the physical world. Combining the physical and virtual instruments resulted in a creation of augmented musical reality.

9. Acknowledgements

The author would like to express his gratitude to ICA-ACROE in Grenoble and Camargo Foundation in Cassis, France for their support.

10. References

[1] Cadoz, Claude, Annie Luciani, and Jean Loup Florenc. "CORDIS-ANIMA: Modeling and Simulation System for Sound and image Synthesis—the General Formalism." Computer Music Journal, Vol. 17, No. 1 (Spring 1993).

[2] Castagne, Nicolas and Claude Cadoz. "GENESIS: A Friendly Musician-Oriented environment for Mass—Interaction Physical Modeling." *Proceedings International Computer Music Conference*, Goteborg, 2002.

[3] Gaver, William W.. "How Do We Hear in the World? Explorations in Ecological Acoustics." *Ecological Psychology*, *5* (4) (1993a), 285 – 313.

[4] Gaver, William W.. "What in the World Do We Hear? An Ecological Approach to Auditory Source Perception." *Ecological Psychology, 5* (1) (1993b.), 1-29.

[5] Dodge, Charles and Thomas A. Jerse. *Computer Music: Synthesis, Composition, and Performance.* New York: Schirmer, 1997.

[6] Puckette, Miller. "Max at Seventeen." *Computer Music Journal*, 26:4 Winter (2002), 31–43.

[7] Schaeffer, Pierre. *Traite des Objets Musicaux.* Paris: Seuil, 1966.

[8] Zicarelli, David. "An Extensible Real-Time Signal Processing Environment for Max." Proceeding of International Computer Music Conference, Ann Arbour, Michigan, 1998.

[9] Zuskinova, Iveta. *Ovciarstvo a Salasnictvo na Liptove (Sheepbreeding and Shepherding in the Litpov Region).* Liptovsky Hradok, Slovakia: Spolok Priatelov Muzea Liptovskej Dediny, 1999.

Reading Perception – Perceiving Literature: an Interdisciplinary Approach

Andreas Lampert* Manfred Nusseck** Jürgen Wertheimer* Heinrich H. Bülthoff**
(*)Universität Tübingen, Comparative Literature, Germany
(**)Max Planck Institute for Biological Cybernetics, Germany
E-mail: andilampe@gmx.de, manfred.nusseck@tuebingen.mpg, juergen.wertheimer@uni-tuebingen.de, heinrich.buelthoff@tuebingen.mpg.de.

Abstract

This paper presents the results of experiments made within the interdisciplinary project between the Department for Comparative Literature at the University of Tübingen and Max Planck Institute for Biological Cybernetics. We examined the following three questions using both psychophysical and structuralistic-hermeneutical methods: a) Are there regularities in the judgments of spatial descriptions by different readers? b) Do readers encode the perceptual perspective of characters during reading? c) Are there correspondences between foregrounding effects and the physiological reaction (galvanic skin response) of readers? The results show that a) that the semantic validation of spatial descriptions showed high homogeneity; b) the method actually showed more about literary strategies concerning object occluding and required the development of new experiment approaches to identify the perspective taken by the reader, that c) the emotional response is quite idiosyncratic but can be roughly divided into two schemes (high or low response).

1. Introduction

In the act of reading, two only apparently prefigured systems interact and influence each other constantly: the text and its reader. Iser [7] used, in his reader response theory, a cybernetic model for this procedure but examined his concept of an 'ideal reader' on a merely theoretical level. We tried in the practical part of our project to analyze aspects of real reader behaviour as suggested by Miall [9]. The detailed descriptions, test arrangements and results will be presented in parts 2 to 4.

The more theoretical portion of our project considered how perceptual phenomena are represented in literature. We claim that literature can focus on and show on a text-immediate level how personal concepts, perceptions, and ideas are transformed into (more or less) communicable signs – words, texts, and images – and how these correspond with or confront historically and/or culturally different forms, norms, and habits of perception. We think that perception (at least as treated in literature) is always a social and dialogical phenomenon since the 'things inside the head' are only accessible and communicable by signs (Eco [3]). Semiotics, however, does not have to look inside the "black box" of human mental representations and their more or less iconographic contents but can instead can focus on how these become functional units in a semiotic process.

We also think that literature plays not only a significant role in the semiotic process in which perceiving is communicated and evaluated, but also in the actual act of perception. Apart from Peirce's assertion that there is also a nonverbal semiotic process (Peirce [10]), it is still unclear if "we only see what we know" or if "we only know what we see". Every act of perception is a hypothetical conjecture based on the current cognitive processing of external facts and on previous cognitions. 'Facts', however, have to be named, so every categorical system is at least partially depended on the words, verbal ideas, and terms that shape and model the world and how it is perceived and described. Besides this problem, literature can not only influence and deform our perceptions, but in extreme cases completely replace one's own perceptions. So literature not only represents models of reality but also can create new possibilities of realities, even if only in the readers' minds.

A literary text may function as an interface that either enhances or hinders the reader's (non-) voluntary intention during and after the reading process to have certain encoded perceptual perspectives (Experiment 1), emotional responses (Experiment 2) or spatial impressions and ratings (Experiment 3) by using special syntactical, semantical, morphological, stylistic and poetical strategies.

2. Experiment 1

Do readers encode the perceptual perspectives of characters during narrative comprehension?

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

2.1. Experiment 1: Hypothesis

Following the design of Horton and Rapp [6], we used literary stories, which described situations in which certain objects were occluded from the protagonists´ point of view. The hypothesis was that *information no longer visible to story's protagonist is less accessible for the reader as well because the reader would assume the protagonist's perspective.* Verification questions about an occluded object should provoke slower reaction times than questions about a still visible object.

2.2 Experiment 1: Design of and Differences to Horton and Rapp

Eleven Participants (6 female, 5 male) with no special literary schooling had to read 15 texts in a randomized order (black lettering on a white background) on a computer screen in a neutral environment.
Each text was shown line per line and participants could continue on to the next line by pressing a key on a computer keyboard. In order to simulate a normal reading experience, the individual lines were centered on the screen, were placed in a typeface similar to that one would find in a book, and were not divided in semantic units but sorted by length (80-90 characters including spaces). All texts were written in an auctorial point of view. In 5 stories, an object was occluded at some point during the story. In 5 other stories, the object was always visible for the protagonist. There were also 5 filler stories for which we asked about objects that didn't occur in the stories. The stories were between 4 and 10 lines long (average 7.4). The object occlusion always happened implicitly (i.e., no story contained a sentence like "X couldn't see Y any more"). After reading a text, participants had to answer a verification question about a particular object. They had to answer the question as quickly and accurately as possible. There was no time limit.
In contrast to Horton and Rapp, we used literary texts that were not specially written for the experiment and differed in form, content, length, object accentuation, and mention (see Section 2.3). All of Horton and Rapp's texts consisted of 7 sentences. In the first three sentences, the protagonist and his environment were introduced. In the fourth, the object is named. In the last two sentences, the object is either occluded or stays visible. Horton and Rapp used a response time limit of 2500 ms.

2.3. Experiment 1: Results and Discussion

We could not confirm the hypothesis of Horton and Rapp. We found no significant differences in reaction time between occluded and visible objects. In fact, participants showed a slight tendency towards slower reactions for visible objects.

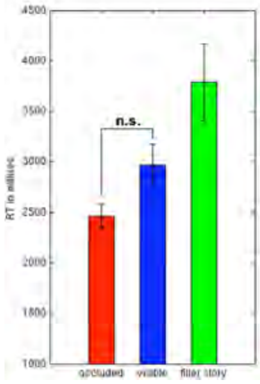


Figure 1: Reaction Times averaged by the three conditions

Our conjecture is that literary texts pursue a somewhat different 'occlusion strategy' than the texts written by Horton and Rapp. Literature often channels the focus of attention on an object despite of, or because of, its occlusion. Texts name objects in different numbers (in our case between 1 to 8 times) and at different points in the story. The concentration of actions and their traceability also vary within a story and texts do not always invite the reader to identify with a protagonist's viewpoint. The texts differ in the emotional allocation of objects and especially their symbolic value and importance for the story.
To have a better control over the input texts, one might change the narration time and the point of view of the stories. Maybe readers identify more strongly with the protagonist's point of view when a story is told by a first person narrator rather than in an auctorial perspective. Such a change avoids the problem that our texts were too heterogeneous and that the Horton and Rapp's texts were too homogeneous. It also allows us to concentrate on the question which literary stylistics could influence the perspective the reader encodes.

3. Experiment 2

Do readers react emotionally to foregrounding effects in literature?

3.1. Experiment 2: Hypothesis

The emotional response to literature is still an underestimated factor researched only in the last two decades. So, we decided following the proposal of Anz [1] to analyze a possible correspondence between

foregrounding effects in literature and the galvanic skin response of readers. The role of emotions in receiving literature has been disregarded since Wimsatt and Beardsley's *The affective fallacy* [2]. They claimed that the emotional effect could not be an objective measure for analyzing literary texts. Wimsatt and Beardsley were not interested in the role of the reader anyway, since they followed the ideas of the *New Criticism* (which focuses exclusively on text-immanent analyses).
Instead of this approach, we had the following hypothesis: R*eaders will show similar galvanic skin responses to certain foregrounding effects in texts.* The term *foregrounding* (see Hakemulder [5]) describes a procedure through which certain text elements get in the reader's focus of attention because they subvert habits and norms of perceptions and/or language systems. This effect of defamiliarzation can be a conscious cognitive process while still affecting the reader's physiology, in our case the galvanic skin response. GSR is, evolutionary seen, a 'fight-or-flight' reaction (Kandel/Schwartz/Jesell [8]), so we supposed that foregrounding effects in literature – as they function as described above – would increase the GSR.

3.2. Experiment 2: Design

The principal design of Experiment 2 was quite similar to Experiment 1. Five participants (3 female, 2 male) had to read 10 texts (randomized order) as described in Section 2.2. The number of read lines was 12 to 41 (average 29.3). Between each pair of texts, there was pause of 60 seconds to avoid 'emotional overlapping' between texts. During the whole experiment, the participants were connected to a galvanic skin response device so we could set the read lines in correspondence to the GSR gradient of each participant. Unfortunately, we could not examine the relationship 'word – gradient'. Further research with eye-tracking devices would be necessary to do this.
The texts were divided into two groups. In the so-called 'neutral' texts, neither feelings nor actions, which could cause a relevant emotional response, were directly described. In contrast, the other texts were the 'affective texts'. Due to shortage of space, we can only present one example text (Figure 2).

3.3. Experiment 2: Results and Discussion

For the Heiner Müller text, reader response is highly idiosyncratic. There are tendencies for a general peak rising in line 1 (where there is a strange conjunction between 'hair of a young woman' and the unusual word – at least in this context - 'to gather together''), line 8 ('bone gripper' and 'cut the chest'), and line 12

1. Das lange Haar der Frau wurde nach dem Wirbel zusammengerafft, der Kopf mit dem Nacken auf
2. den Holzblock gelegt, (die) Stirnhaut parallel zu den Augenbrauen durch(ge)trennt, nach
3. hinten abgelöst, mit (der) Knochensäge (das) Schädeldach abgetrennt (leises Knirschen wie bei
4. einer Laubsägearbeit), Schädel nach hinten geklappt, mit den Händen das Gehirn aus einer
5. Einbettung gelockert, Organ auf Glasplatte gelegt. (Sektionsbericht) Die ersten Schnitte
6. parallel oberhalb der Schlüsselbeine, dann von der Halsgrube abwärts die Haut über dem
7. Brustbein aufgeschlitzt, rechts am Nabel herum und weiter abwärts über dem Bauch. Mit der
8. Knochenzange den Brustkasten von beiden Seiten her schräg zum Schlüsselbein durchschneiden,
1. Brustbein mit beiderseits anschließenden Rippenstücken in der Mitte hochklappen
2. (Sektionsbericht 2) Während einer Operettenaufführung (Lehar: Land des Lächelns) die
3. Zwangsvorstellung, daß ich zwei Nadeln in die (wahrscheinlich blauen) Pupillen der Soubrette
4. stecke. Wann die Nadeln herausgezogen werden, laufen die Augen aus. (Tagebuchnotiz)

Aus: Heiner Muller: Die Prosa. Frankfurt 1999. S. 165

Figure 2: Text by Heiner Müller. Foregrounded elements are marked red.

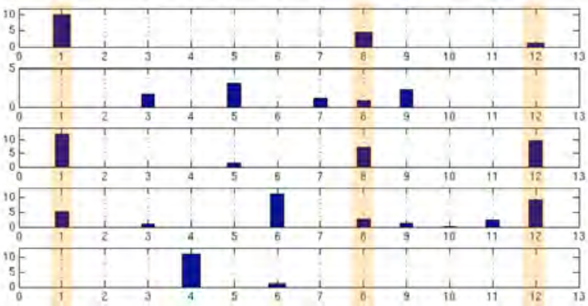


Figure 3: significant peaks per line of all readers

(here we have the shift from auctorial point of view to first person narration and an abrupt action of cruelty: 'to put needles into eyes'). There is also the fact that participants showed either constantly high (participants 1 and 5) or low response (participants 2 and 4). In future experiments, we definitely need a) more participants to get more objective results and b) stronger contrasting texts (e.g., non-literary texts or political speeches). It also would be interesting to analyze the relationship between personal appraisement of emotion and actual GSR results.

4. Experiment 3

Do readers respond to spatial descriptions in literature with similar semantical judgements?

4.1. Experiment 3: Hypothesis

Reading spatial descriptions in a story evokes mental representation in readers. We wanted to analyze (following the experiment of Franz [4]) *if different readers would judge narrated rooms in a similar way.*

4.2. Experiment 3: Design

As in the other experiments, 21 participants (14 females, 7 males) had to read 10 texts (randomized order; 7 to 42 lines long, average 14.1). After each text, participants had to rate the described rooms using 5 semantic differentials, the order of which was randomized. The semantic differentials had a seven point scale and were as following:

1. hässlich – schön (ugly – nice)
2. geschlossen – offen (closed – open)
3. langweilig – interessant (boring – interesting)
4. unangenehm – angenehm (unpleasant – pleasant)
5. ungewöhnlich – gewöhnlich (unusual – usual)

We chose texts written in an auctorial perspective in which no comments by intermediatory story protagonists could influence the room description. We also tried to avoid any other prejudicial text-immanent elements.

4.3. Experiment 3: Results and Discussion

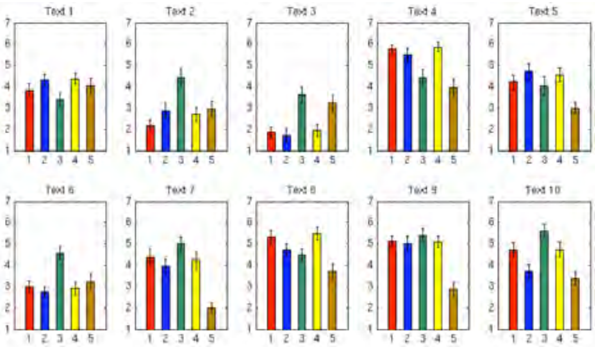


Figure 4: Ratings of the participants for each room. X-axis represents the 5 semantic differentials; Y-axis: ascending sequence from negative to positive judgements.

We found a correlation between 'ugly', 'closed', and 'unpleasant', but no correlation between 'interesting' and either 'ugly' or 'nice'. There was also a correlation between 'unusual' and 'interesting', so the last both results showed that readers were fascinated more through aesthetic form than content. Small and narrow rooms produced rather negative responses (text 3). Interestingly, text 3 was also the only one in which a protagonist (and his bad mood) could have influenced the room impression. The sex of the participants had no significant impact on the pattern of results. Future experiments could concentrate on culturally different reactions to room descriptions, if there is a different reaction on old and new texts, and if it is significant if there are actions described in a room or not or the room is seen by a protagonist or not.

5. General Discussion

The interdisciplinary approach showed new perspectives and insights on perception and more research needs to be done. Where psychophysical experiments help to understand the emotional and cognitive process of re-/perceiving literature, semiotic literary sciences can analyze in which historical, ethical and social contexts individual or collective perceptions are made and transformed into signs. Both the brain and literature are so to say filter-tools that form and construct 'as-if-realities', but literary texts boost ambiguities and polysemies where in contrast the biological perception-apparatus is inclined to establish a monological and unambigous reality.

References

[1] Anz, Thomas. Plädoyer für eine kulturwissen-schaftliche Emotionsforschung.", literaturkritik.de Nr. 2/3, 1999 (www.literaturkritik.de/public/rezension.php?rez_id=47).

[2] Beardsley, William/Wimsatt, Monroe, „The affective fallacy", The Verbal Icon: Studies in the Meaning of Poetry. P. 21-39, Lexington: University of Kentucky Press, 1954.

[3] Eco, Umberto. „Kant und das Schnabeltier", DTV, München 2000.

[4] Franz, Gerald/Heyde, Markus v. d./Bülthoff, Heinrich H.,"An empirical approach to the experience of architectural space in virtual reality", Automation in Construction 14 ***(2)***, 165-172, 2005.

[5] Hakelmulder, Willie van/Peer, Frank, " The Pergamon Encyclopaedia of Language and Linguisitics", pp. 546-551, ed. Keith Brown. Oxford: Elsevier, 2006, vol. 4,.)

[6] Horton, William/Rapp, David, "Out of sight, out of mind: Occlusion and the accessibility of information in narrative comprehension." Psychosomatic Bulletin & Review 10 (1), p. 104-110, 2003.

[7] Iser, Wolfgang. „Die Wirklichkeit der Fiktion", p. 277-325, ed. Warning, Rainer, Rezeptionsästhetik, München, UTB, 1994.

[8] Kandel, Eric/Schwartz, James/Jesell, Thomas, „Principles of neural science", 4th ed. McGraw-Hill, New York.

[9] Miall, David, "On the necessity of empirical studies of literary reading", Frame. Utrecht Journal of literary reading 14, 2-3, p. 43-59. 2000.

[10] Peirce, Charles S., „Writings", Indiana, University Press, 1982-83.

Ergotic Sounds
A new way to improve Playability, Believability and Presence of Digital Musical Instruments

Annie Luciani, Jean-Loup Florens, Damien Couroussé, Claude Cadoz
ACROE, INPG, 46 avenue Félix Viallet, 38000 Grenoble, France
E-mail : Annie.Luciani@imag.fr, Jean-Loup.Florens@imag.fr, Damien.Courousse@imag.fr, Claude.Cadoz@imag.fr

Abstract

We explore here is how "ergotic gestural-sound situation", i.e. the situation in which the instrumental gesture and the produced sound are intimately energetically linked, is a way to guaranty the playability, the believability and the presence of the musical digital instrument. The chosen experimental situation is the evaluation of a musical pattern on a bowed string performance that consists in maintaining the continuity of the sound when changing the bow direction. We present some surprising results, thanks to the hiqh quality real time "cello-like" haptic simulation, implemented on the ERGOS haptic technology.

1. Ergotic sounds

In the musical performance situation, two types of relation between the human gestures and the produced sound are usually distinguished:

• Non-instrumental musical practices, like the musical conductor situation, or the control of synthetic sounds parameters by mapping techniques [1][2].

• The instrumental musical experience, as when an instrumentalist is playing a physical instrument.

In non-instrumental practices, there is no energetic relationship between the human gestures and the sound so-controlled. Conversely, in the instrumental musical experience, the performer and the played object are physically dynamically coupled during the playing [3]. The produced sound engraves the physical energy exchanged between the performer and the physical musical instrument. It is what we call "ergotic sounds" to relate about an « ergotic relation to the sound » according to the typology of relations between human and environment proposed by C. Cadoz [4][5].

As emphazised by O'Modhrain et al. [6], we assume that the instrumental musical experience is an emblematic case of enaction, exhibiting the main features of the enactive theory of cognition:

• The world without representation [7]: the representation of the instrumental situation is only the situation itself.

• Cognitive categories (as here for example "musicality) do not pre-exist: they are emerging from the interaction with the environment, i.e. from the instrumental situation itself.

• Enactive knowledge acquired during the experiment is robust: the learning of the task (here is the performance) and the cognitive category (here is the musicality), that are emerging during the performance itself leads to a robust know-how.

The role of the energetic exchange between the instrumentalist and her instrument is well recognized when playing mechanical musical instruments. We may assume that the ergotic relation to the sound is an important feature for the playability of the musical instrument, that have to be re-introduced in digital musical instruments to improve their playability, their believability, the feeling of their sensory presence and of their presence in hand.

However, it is not easy to experiment such assumption in real mechanical situation and digital musical instruments can be very new experimental set-up to understand what it happens during an ergotic instrumental situation.

In the following, we describe the simulator we designed to catch the main feature of ergotic musical situation, and a first set of results obtained from a pilot experiment performed with 10 subjects.

2. The Virtual Cello-like Platform

2.1. The Physical model of the bowed string

Current implementations of interactive sounds are usually based on the mapping concept. Indeed, in the mapping concept, there is no energetic consistency between the process that produces the sound and the process that maps the gestural data acquired through sensors to the parameters of the sound synthesis process. The data flow between both is unidirectional, from the gesture to the sound.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

When extended with force feedback interaction, as done in [6][1][8][9], the gesture side is improved by a local physical model that produces the force sent to the hand, but there is no longer retroactive interaction from the vibrating string to the manipulation part. Even though we may assume that the vibration of the string is not felt by hand, it is known in acoustics that the interaction from the string to the bow at the acoustical frequency plays an important role in the bowing.

A virtual bowed string model, that respects the bilateral physical interaction along all the instrumental chain, between the hand and the bow and between the bow and the vibrating structure, has been implemented in [10]. It is composed of two sets of interactions (Figure 1):

• two bilateral interactions between the bow and the vibrating string: (1) the buffer interaction for the transversal motion representing the collision and the pressure between the bow and the string and (2) the friction interaction for the lateral motion.

• A bilateral interaction between the bow and the hand via a 2D force feedback device returning the pressure and the friction forces from the bow to the hand.

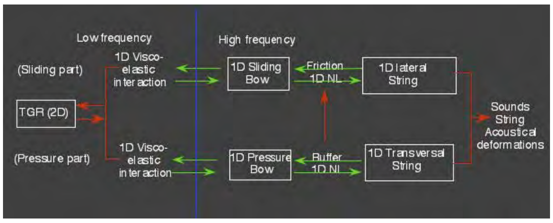


Figure 1.Ergotic Physical model of the bowed string

This model guaranties the “ergoticity of the situation” as it is the case in real mechanical instrumental playing; there is no break in the rendering of the energetic consistency between the hand and the hear.

2.2. The high quality ERGOS Experimental Platform

The technical platform has to implement a real time simulation of the previous model with a sufficient quality to be sure that the main features existing in the real situation could be caught.

Florens and his co-workers [10][11][12] developed the most reactive implementation based on a synchronous computer architecture and a high fidelity haptic ERGOS technology. In this implementation, the previous bowed-string model has been implemented with a two overlapped frequency loops: 3 KHz for the hand-bow loop and 44 KHz for the bow-string loop. An impressive quality of the sounds has been reached, encouraging to go further in the understanding of the sensitivity of the string by the hand, by asking the question: Is the gestural sensitivity of the vibrating string, although filtered by the bow and by the mechanical human body, important, and for what? Such question has never been experimented neither in real life situation nor in virtual reality based situation, although the “feeling” of the vibration of the string in the fingers is often noticed by instrumentalists.

To elicit the role of the “ergoticity” in the playing – i.e. the role of the energetic link between hands and sounds, one way is to implement an experimental virtual reality platform able to render the feeling of the vibrating string within the hand. Consequently all the part of the physical previous model must be implemented at the frame rate of the sound, i.e. at 44 KHz (Figure 2).

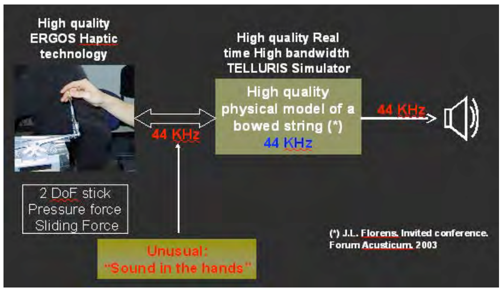


Figure 2. Functional diagram of the high quality 44 KHz experimental ERGOS platform

To do that, the entire model has been implemented on a DSP board directly connected to the force feedback ERGOS device by high reactive DA/AD converters and the sound is directly picked onto the DSP board.

The reactivity between input action and force feedback as well as between input action and sound are both of 1/44000 ms, and totally synchronous at that rate.

2.3. The experimental protocol

We select a feature that is considered as difficult to perform and to learn in real practices. It consists in maintaining the continuity of the sound when changing the bow direction.

This feature is subjectively well - identifiable by the performer and by the audience.

But it is also objectively observable on the signals as shown in the snapshot on Figure 3. When changing the direction of the bow by maintaining the continuity of the sound, the bowing movement is inverted (3), but the phase of the string vibration is not changed.

Performer and audience are invited to point out when the goal is reached and the moment when the subjective success occurs.

The experiment has been performed 10 subjects: 25 to 55 years old, 4 females and 6 males, 6 non professional musicians (3 totally novices and 3 with a slight instrumentalist expertise) and 4 musicians. One of them is a native blind people. The audience is composed of the previous subjects and 10 others persons, most of them being not aware of the activity of the laboratory.

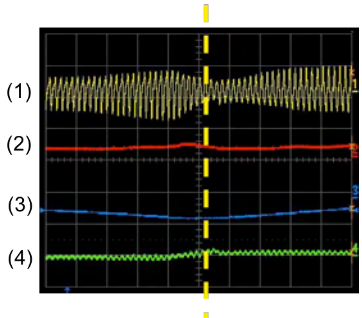


Figure 3. Signals corresponding the bow direction changing. (1) Acoustical vibration of the string, (2) pressure force of the bow on the string, (3) Displacement of the bow, (4) Sliding force of the bow

The subjects play freely the duration they want.

Subjects are confronted to four types of gestural feedbacks: (a): no friction; (b): slight friction, i.e. lighter than in the real case; (c) normal friction, i.e. closer to the real case, (d) exaggerated friction, very higher than in the real case.

Success is observed on objective signals and by asking the performer and the audience.

3. Results

3.2.1. Occurrences of the success of the task

The main observations are the following:

• With Slight or Normal Friction: Most of the people reach several times the goal after a very short time of trial (no more than 15 minutes), even if they estimate they are bad or non expert. The occurrence of success is greater than in real case.

• With Exaggerated friction: A very few number of people reach the goal. The best scores are obtained by the blind people. For all of them, the playing is difficult and non satisfying.

• With Null friction: A very few number of people reach the goal. The best scores are performed by the blind people. All the scores are better than with exaggerated friction. However, most people estimate that they can reach better after more learning.

We may conclude that adequate well-tuned ergotic relation to the sound is important to reach this goal.

3.2.2. Cognitive styles

A very first observation was that the ways of exploration depend on a priori cognitive styles. During the first minutes of playing, some people are caught by non dynamic features such as spatial features (geometrical properties as curvatures) or musical features (timbre or pitch). But progressively, and quite quickly (less than 15 minutes), all are attracted by the dynamic of the playing and started to explore it with new non predetermined ways: changing the force, the accelerations, the velocities, the trajectories, etc.

3.2.3. Modes of Playing and dynamic adaptation

A very obvious observation is made by all the persons of the audience and confirmed when examining the video movies: there is a continuous adaptation of gestures to find the best way of manipulation to reach the goal. Some examples are:

• Exploration of various modes of grasping and postures: With fingers, hand palm, deployed arm, strong full hand grasping, etc.

• a wide exploration of dynamic strategies: bow direction changing by soft round turning, road turns, Möbius-style movement, elliptic trajectories, modulating the cinematic of the gestures, acceleration/decceleration near the point of changing, relaxing the pressure before or after the turn point, etc. (see photographs on Figure 4).

Figure 4. Exploration of different ways of playing to reach the goal.

3.2.4. Continuous Dynamic Learning

Despite the morphological non - similarity with a real instrument (the displacement of the bow is of about some centimeters!), all the people learn very fast how to perform the playing. After less than 15 minutes, all the people were at ease with the instrument, improving very quickly the quality of their gestures.

Very quickly also, they start free exploration of a wide range of dynamic strategies to reach the goal. They declared that imagined a priori strategies are not correct. They learn on the fly “to be within the situation” and so doing they find very original strategy to reach the goal such as: *“Relax and let the bow act by itself just before the turn”* (Quotation of most of the people)

The best were the blind people for whom it is confirmed that she is a fine expert in haptic - audio strategies and, surprisingly, people that never manipulate real instruments or that do not have non predetermined cognitive styles.

3.2.5. Playability, creativity and Presence

For all the performers, “Exaggerated friction” is non affordant. For example, persons starting with this case do not understand the task itself. However, we observe that this non-affordance stimulates the discovery of new efficient ways of manipulation and of playing, leading to create new type of gestures and sounds. Affordance and creativity seems different concepts that can be sometimes and perhaps contradictory.

Obviously, the normal friction is the most playable and pleasant for all the performers. It increases to duration of playing stimulating creativity but within a more conventional attitude with less original exploration.

But above all, the unexpected result we are particularly proud, that is very new and very promising, relies on the spontaneous remark made by all the performers happily surprised by the “strong presence of the string in hand”. Thanks to the 44Khz audio-haptic simulation, a strong and unanimous of “*The string Presence*”, “*the string in the fingers*”, “*the string is really here*”, etc. Something of new appears, comparing with our first high quality implementation of ergotic sounds in which the haptic parts were running at 3KHz only. This could explain the strength and the predominance of the sound in the identification of the object by a co-reinforcement of the haptic and audio feedbacks in the sense that the acoustical vibration of the string and the “feeling of the string in hand by the haptic feeling of its vibration” convinces reliability of the indubitable presence of “that instrument”.

4. Conclusions

The experiments presented here were motivated by an assumption that is well-accepted in the mechanical musical situation but quite impossible to evaluate without a specific virtual instrumental playing implementation: examine if ergotic audio-haptic situation - i.e. a situation in which the physicality of the interaction is maintained within the whole instrumental chain between the hand and the hear, the gesture and the sound - plays a core role in performing complex and subtle musical cues. The first set of performed experiments presented here are promising. Adapted and well-tuned ergotic sounds situation enhances instrumental learning and playability. It allows the performer to dynamically adapt her manipulation “on the fly” for the success of the goal. It supports very fast instrumental learning through very quickly acquired exploration strategy of manipulation. But above all, the most unexpected and promising result relies on the spontaneous remark made by all the performers happily surprised by the “strong presence of the string in hand”, triggering a strong feeling of presence of the string, thanks to the never realized 44Khz audio-haptic simulation, implementing really what we called at the beginning of the research “ergotic sounds”.

References

[1] Nichols C: “the vBow: a Virtual Violin Bow Controller for Mapping Gesture to Synthesis with Haptic Feedback” – in Organised Sounds – Leigh Landy ed. - Leicester, United Kingdom, 2002.

[2] Andy Hunt, Marcelo M. Wanderley, and Matt Paradis. 2003. "The Importance of Parameter Mapping in Electronic Instrument Design." Journal of New Music Research, 32(4):429-440.

[3] Essl G. and O’Modhrain S.. « An enactive approach to the design of new tangible musical isntruments ». Orgaised soubnds 11 (3). 285-296. 2006. Cambridge University Press.

[4] Cadoz C. "Instrumental Gesture and Musical Composition". Proceedings of International Computer Music Conference - Cologne 1988.

[5] Cadoz C., Wanderley M. (2000). “Gesture-Music” – in “Trends in Gestural Control of Music”, M. Wanderley and M. Battier, eds, pp. 71-94 -2000

[6] O'Modhrain S., Chafe C.: “Incorporating Haptic Feedback into Interfaces for Music Applications” – in proceedings of ISORA, World Automation Conference, 2000.

[7] Clark Andy and Toribio Josepfa. “Doing without representating?” Synthese 101:1994 401-431. Kluwer Academic Publishers. 1994.

[8] Rimell S, Howard D M., Tyrrell A M., Kirk R, Hunt A – “Cymatic: Restoring the Physical Manifestation of Digital Sound using Haptic Interfaces to Control a New Computer Based Musical Instrument” – International Computer Music Conference, Sweden, 2002

[9] Sinclair Stephen, Wanderley Marcelo M. « Defining a control standard for easily integrating haptic virtual environments with existing audio/visual systems ». Proceedings of NIME 07. New York, NY, USA.

[10] Florens Jean-Loup. “Real time Bowed String Synthesis with Force Feedback Gesture”. Invited paper. Forum Acousticum. Sevilla. November 2002.

[11] Florens JL, Luciani A, Castagne N, Cadoz C. “ERGOS : a Multi-degrees of Freedom and Versatile Forcefeedback panoply”. Proceedings of Eurohpatics, pp356-360, Germany, 2004.

[12] Luciani Annie, Florens Jean-Loup, Castagné Nicolas. From Action to Sound: a Challenging Perspective for Haptics. Proceedings of WHC 2005.

Acknowledgements

This research work has been performed with the support of the French Ministry of Culture and the FP6 NoE Enactives Interfaces.

Non-Conscious Control of Sound Spatialization

Mark T. Marshall   Joseph Malloch   Marcelo M. Wanderley
Input Devices and Music Interaction Laboratory, and
Centre for Interdisciplinary Research in Music Media and Technology,
McGill University, Montreal, QC, Canada
E-mail: {mark.marshall, joseph.malloch, marcelo.wanderley}@mcgill.ca

Abstract

This paper presents ongoing work into the design of interfaces for live musical performance using gesture controlled sound spatialization. In particular, we discuss the use of performer's gestures to allow them to manipulate spatialization parameters without their conscious control. A number of motion capture sessions involving a cello performer are described through which we determined a set of gesture parameters which are useful for this purpose. We also present our implementation of a wireless sensor system to allow control of spatialization parameters using cello performer gestures. Finally, we provide some guidelines for succesfully mapping gesture control to sound spatialization parameters.

1 Introduction

Musical performance intrinsically encompasses spatial elements, in the arrangement and directivity of musicians and instruments, the acoustics of performance and listening spaces, and the relative positioning of audience and performer. Sound spatialization effects have long been used by traditional composers and orchestras, but it is only since the advent of electric sound reinforcement and diffusion that the bulk of spatial parameters may be changed dynamically or be controlled in real-time. For a detailed review of the history of spatial music see [7].

As part of a project on the compositional uses of gesture controlled sound spatialization, we have been investigating possible performer roles and developing systems to explore the possibilities of this type of control [3]. The main aim of this work is to investigate and develop interfaces which can provide rich, intuitive control of compositionally interesting parameters in a sound spatialization system, and are above all compatible with expert musical performance.

Previously we have discussed the performance and mapping issues involved in implementing gesture control of spatialization parameters[2], with especial focus on resolution (of gesture, sensing, mapping, and subsequent control), integrality and separability, current vs. ballistic control strategies, and the *cognitive load* carried by a performing instrumental musician. This last issue addresses the concern that not all performers have spare attention for controlling something other than their instrument [1]. Taking this into account, we have divided performer control into three main approaches, one of which is the control of spatialization parameters by a performer's gestures *without their conscious control*.

If conscious control is desired, gestures must be chosen such that they can be performed without disturbing the instrumental performance, and it is assumed that the performer has spare attention for this task. For non-conscious control, the mapping relationships between performer movement and spatialization effect becomes an indirect compositional process rather than instrument augmentation or performer interpretation. Rather than asking the performer to deliberately manipulate spatial parameters, the composer or designer must plan instrumental movement with thought to the spatial effect as well as acoustic sound production.

This paper will focus on the third approach, specifically on explorations into non-conscious control of spatialization parameters using the movements of a performing cellist. An explanation of the process used to choose which movements to use for control will be given, followed by steps taken to move the chosen control systems from the lab to the concert hall. Finally, the implications of this type of control for composers, performers, and audience will be discussed, with insights into how to approach mapping gesture control to sound spatialization parameters successfully.

2 Aquiring Performer Gesture Data

In order to inform the choice of control strategies, it was necessary for us to make a number of recordings of performer gestures which could be analysed to extract useful data. To this end, we performed a number of motion capture recording sessions using various technologies. These sessions focussed on recording and analysis of cello performer gestures.

2.1 Preliminary Motion Capture

The first capture session made use of a VICON motion capture system[1]. Markers were placed on the performer's body, concentrating on the upper body. The main goal of this recording was to identify movements useful for controlling spatialization which could later be captured using a more basic system suited for performance use. It was expected that the most useful data would be related to the bowing arm, but it was hoped that other useful data would also be discovered.

For this session, we recorded a number of performances with a single player. To allow us to examine a variety of material, she performed a number of simple scale excercises, followed by performances of 2 different pieces, each of which had a very different tempo.

As expected, we found that we could extract bowing data from the movements of the bowing arm in the axis perpendicular to the strings. This could give us information on bowing speed, position and energy. However, further analysis of the data indicated that the performer's center of mass might also prove a useful parameter. The recordings revealed a regular low-frequency oscillation in the performer's center of mass, the frequency of which appeared to be related to the piece being played. Figure 1 shows movement of the performer's center of mass in the forward/backward direction for the 2 recorded pieces. The second piece, which is faster than the first, shows a wider range of movement, more deviation from the neutral position and has additional high frequency movements not present in the recording of the slower piece. We also noticed a tendency for the performer to lean forward at the beginnings of musical phrases. Each of these artifacts might prove useful as controls in performance situations.

2.2 Subsequent Capture Sessions

Having determined that we could extract some useful data from the movement of the bowing arm, we began designing a system to allow us to record this data that could also be used in live performance. We decided to make use of accelerometers, as these are unobtrusive enough for performance use, but can be used to acquire both acceleration and rotation data.

Our design made use of ST Microelectronics LIS3L02AS4 3-axis accelerometers which were worn on the performer's forearms. These accelerometers were connected to the computer using a Teabox sensor interface from Electrotap[2], which recorded the accelerometer

[1] http://www.vicon.com/

[2] http://www.electrotap.com/

Figure 2: The accelerometer-based capture system in use. The accelerometer board is being worn on the performers bowing arm.

data at a sampling rate of 1 kHz and a resolution of 16 bits. This provided high-quality data for our analysis.

Using this system, we recorded a large number of cello performance techniques, including different styles and speeds of bowing, pizzicato and extended techniques such as bowing on the bridge and on the body of the instrument. Figure 2 shows this system in use during the capture session.

From this session, we determined that the accelerometers could be used to measure most of the bowing parameters already identified. In addition to the bowing speed and energy, we found that we could determine which string was being bowed, from the angle of the bowing arm relative to the instrument. We also found it possible to determine a measure of the left-hand position on the fingerboard, using the angle of the fingering arm relative to the vertical plane.

A final motion capture session also took place, this time using 3-axis accelerometers together with a Bluetooth interface. This allowed wireless transmission of the data, which offers a distinct advantage for the performer, but at a slower update rate. In this final session, a number of recordings were made which measured the identified gestures to allow us to determine if the 200Hz update rate of the wireless sensors was sufficient for our purposes.

2.3 Features

Overall from these motion capture sessions we have discovered a number of gesture parameters which can be easily sensed and which may be useful for controlling spatialization. These parameters are:

- relative position of playing on the fingerboard, extracted from left arm rotation data

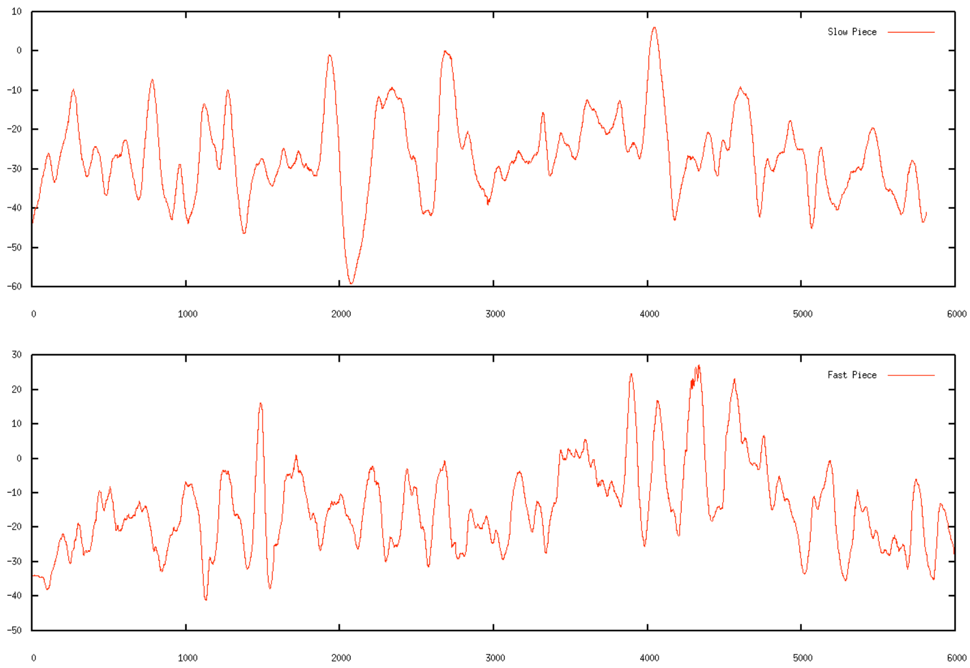


Figure 1: Center of mass movements in the forward/backward direction for the 2 recorded pieces. (a) Low tempo piece and (b) high tempo piece. Note the extra high frequency movements in the faster piece.

- current string being bowed, measured from bowing-arm rotation data
- overall bowing energy
- performer center of mass

It may also be possible to extract more data from the performer's gestures, including the recognition of certain playing techniques based on combinations of the other data.

3 Moving Towards Performance

In order to make use of these gestures in performance we needed to develop reliable, robust methods of sensing the necessary movements. With this in mind and based on the results of our motion capture sessions, we implemented a system which measures acceleration and rotation data from which we can extract the required bowing and fingering gesture parameters. The final performance system makes use of an Xsens Technologies Xbus kit wireless motion capture system, which provides drift-free calibrated acceleration and orientation data.

The data from the wireless sensors is read using a patch in the Max/MSP graphical programming language, which performs recording and analysis of the incoming signals. Signals are recorded into multichannel audio files which allow for easy processing in software such as Matlab. It should also be noted that this patch offers the possiblity of loading previously saved gesture data to allow for simulations based on recordings of real gestures.

Currently, the patch extracts bowing speed, bowing energy, the number of the string being bowed, and left-hand position on the fingerboard, but is easily expandable for other data. Communication between the patch and the spatialization system occurs using OpenSound-Control (OSC) messages[6] which are sent over the network. This allows us to peform the gesture capture on one computer, while running the spatialization on another, thus spreading the processing load across machines.

For the performer center of mass measurement, we decided to build a pressure-sensitive floor. A 1m by 1m plywood floor was built with 4 evenly-spaced force sensing resistors (FSRs) attached to the underside. The outputs from these FSRs are converted to digital values and sent over a USB connection to the computer. By measuring the difference in signal across the FSRs it is possible to estimate the center of mass of a performer standing on this floor.

We initially tested this floor with a number of different users in standing positions and the output proved reliable. However, some complications arise in the case of a seated performer (such as our cello player). Depending on the type of chair used by the performer, the performer's mass can be distributed more or less evenly. This requires some calibration of the system to allow for such differences. Another possible option would be to embed the FSRs in the chair, rather than on the floor underneath it, as described in [4].

Data is read using a patch in Max/MSP, which incorporates both a visual and numeric representation of the center of mass estimated from the pressure sensors. We once again use OSC for the communication between this patch and the spatialization system.

4 Discussion

The ultimate aim of this work is practical: to provide knowledge and tools for composers who wish to use gesture control of spatial effects within their pieces. The solutions and systems are not intended to directly shape the aesthetic effect of the works, but rather to enable composers to produce the acoustic and performative impact they desire. Without presuming to label approaches to an artistic work as "right" or "wrong," our experiences with multiple composers, performers, systems and venues have yielded suggested guidelines which may minimize performer distraction and maximize mapping success when making use of non-conscious performer control of spatialization parameters.

Map resting body states to neutral spatialization states. If resting body states are mapped to dramatic or obvious spatialization effects, the performer is likely to become distressed, as they will be unable to "stop" performing. The field of digital musical instruments has shown that the ability to *stop* sounds and processes is essential for a feeling of being in control[5].

Use non-conscious control for high-level spatialization control. High-level spatialization parameters, such as system changeability, the flocking behaviours of clouds of sound sources, or control over other algorithmic behaviours, may be successfully controlled using ancilliary performer movements. If the mapping relationships are obscure enough, the performer will be less likely to be tempted to try to "take control."

Use non-conscious control for subtle spatialization control. Similarly, subtle effects, such as sound spread or diffusion in the space or boundary reflectivity, may be less distracting for the performer than direct source position or volume.

Avoid changing mapping relationships dynamically. If the audience is intended or permitted to perceive the nature of gesture control, this effect may be confused or destroyed by changing scaling or transfer functions. In the case of very subtle effects this may not be an issue, since it is unlikely that the audience will understand the mapping relationship.

5 Conclusion

This paper described an approach to the use of performance gestures for controlling parameters of a sound spatialization system. Using motion capture systems we have identified parameters of a cello performer's gestures which can be detected and used for such control. Testing has indicated that such control can prove useful to allow interactive sound spatialization without placing much additional load on the performer. There now exists the opportunity to make use of this system in live performance and a number of such performances are planned, beginning with a performance as part of Enactive/07.

6 Acknowledgements

This work has taken place with the support of the *Natural Science and Engineering Research Council of Canada* and the *Canada Council for the Arts*. Thanks to Chloé Dominguez for taking part in the motion capture sessions and to Delphine Bernardin for running the initial session. Thanks also to Sean Ferguson and Georgios Marentakis for discussions and comments regarding this work.

References

[1] P. Cook. Principles for designing computer music controllers. In *Workshop in New Interfaces for Musical Expression*, 2001.

[2] M. T. Marshall, J. Malloch, and M. M. Wanderley. Gesture control of sound spatialization for live musical performance. In *Proc. of the 7th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW2007)*, Lisbon, Portugal, 2007.

[3] M. T. Marshall, N. Peters, A. R. Jensenius, J. Boissinot, M. M. Wanderley, and J. Braasch. On the development of a system for the gesture control of spatialization. In *Proceedings of the International Computer Music Conference*, pages 260–266, 2006.

[4] L. McElligott, E. Dixon, and M. Dillon. 'peglegs in music' - processing the effort generated by levels of expressive gesturing in music. In *Proc. of the 2002 International Conference on New Interfaces for Musical Expression (NIME02)*, pages 6–10, Dublin, Ireland, 2002.

[5] D. Wessel and M. Wright. Problems and prospects for intimate musical control of computers. *Computer Music Journal*, 26(3):11–22, Fall 2002.

[6] M. Wright, A. Freed, and A. Momeni. Opensound control: State of the art 2003. In *Proceedings of the 2003 Conference on New Interfaces for Musical Expression*, 2003.

[7] R. Zvonar. A history of spatial music, 2006. Available online at http://cec.concordia.ca/econtact/Multichannel/spatial_music.html.

Validating kinematic displays for the perception of musical performance

Manfred Nusseck[*] Marcelo M. Wanderley[‡] Erwin Schoonderwaldt[‡]
(*) Max Planck Institute for biological cybernetics, Tübingen, Germany
(‡) IDMIL, McGill University, Montreal, Canada
E-mail: manfred.nusseck@tuebingen.mpg.de, marcelo.wanderley@mcgill.ca, schoonderw@music.mcgill.ca.

Abstract

Human gestures contain certain characteristics and meanings in communication and represent a link between intention and body. This paper describes a pilot study investigating the role of ancillary musical gestures in understanding musical meaning from the listener's standpoint. We conducted a perceptual experiment using motion-capture recordings of musicians. Participants were presented video recordings and reconstructed point-light displays of music performances. By asking them to rate certain music-related parameters we found that abstract motions of the point-light displays yielded similar ratings to those of the real recordings. This suggests that pure body motion seems to be sufficient to communicate certain musical impressions.

1. Introduction

Playing musical instruments is an interaction of both sound and body movements. These movements are tightly linked to each other so that ancillary performance gestures seem to have an intrinsic relationship to the musical performance. Why are these gestures performed? What do they communicate to the audience? How important are body motions for the music experience of the audience? In this study, we examined to which extent visual aspects contribute to the communication process between performer and listener and how they influence perceptual and aesthetic judgments of music. This paper describes a pilot-study investigating whether pure body motion supports similar aspects of musical performance as a real video. It validates whether or not kinematic displays of musicians can be sufficient representatives in judging certain musical parameters. The study is part of a larger project addressing the general role of body motions in perceiving musical performances.

Gestures in music are the topic of a large field of research. Previous studies proposed a general categorization into (a) gestures that are involved in playing the instrument and are related directly to the production of the sound, and (b) gestures that are part of the performance but not produced for the purpose of sound generation [1] [11]. The latter are so-called "ancillary gestures" and are considered as being connected to the expressiveness of the performance [5]. These gestures contain body sways and subtle facial motions, such as eyebrow movements or facial expressions. These visual cues have been found to influence musical intelligibility [9].

Investigating performances of clarinet players, Wanderley et al. [11] showed that ancillary gestures occur frequently in musical performances, even though they are not essential for it. Furthermore, they varied considerably across performers, but are more consistent within performers. The findings showed that these gestures are not randomly performed, but more an integral part of the performance process.

During learning of a new musical piece, Chaffin and Logan [2] showed that musicians' performance cues move through stages from basic and effective to expressive, in a more holistic sense. Body movements, however, are also related to performance interpretation. Dahl et al. [4] showed that video sequences of marimba players, where only the upper body was presented, clearly support emotional intentions. With a multi-modal design in which participants either saw, heard, or both saw and heard musicians performing, Vines et al. [10] found that visual perception of musical performance interacts with auditory perception. They also showed that ancillary gestures are linked to structural features of the music, i.e. phrase boundaries or extended notes.

Given that live performances or even video recordings of musicians contain a large amount of contextual and situational information, such as the players clothes, hair style, body shape, or the environmental background (high-level cues), it is difficult to know what influences a perceiver's judgement. To eliminate these high-level information sources, research started to use so-called "point-light displays", adapted by Johansson [6]. With motion-capture systems, the positions of markers, attached to major body joints on the player, are tracked during the performance. If displaying these positions as dots on a screen, it creates a stimulus based solely on the spatio-temporal movements of body and limbs. Therefore, point-light displays separate performance gestures from all high-level cues so that only the body motion is still present. Previous studies have shown that for human motion recognition point-light display raised similar results as natural videos [8].

Using this technique, Davidson [5] recorded violin players performing three different levels of expression: deadpan, standard, and exaggerated. In a perceptual experiment to recognize expressive intentions, she found that the different expressions were clearly perceivable from point-light displays alone.

Additionally, Davidson [5] found that by presenting the stimuli either visual only, sound only, or multimodal, the gesture of the musicians gave a better indication of expressive intent than the sound. This showed that visual cues seem to have a strong influence on music perception.

In this study we conducted a perceptual experiment investigating if pure body motions can be sufficient to allow rating of music-related parameters. Since the visual aspects seem to be an important source of information, we compared ratings of point-light displays with real videos.

2. Experiment

2.1. Method

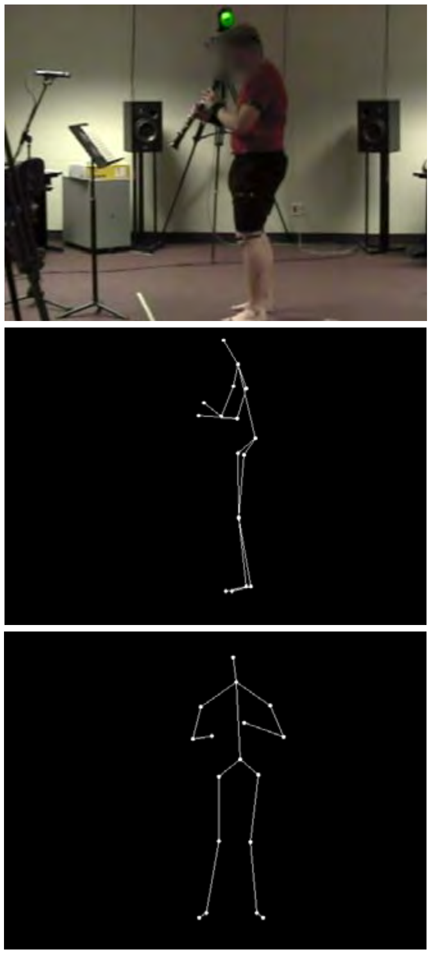
2.1.1. Stimuli

To create the set of stimuli we used motion-capture data of four different clarinet players (3 male, 1 female) performing Brahms Sonata 1, Op 120. The recordings were done at the IDMIL Lab of McGill University. In addition to the motion-capture, the players were recorded using a camcorder. All players stood in front of a music stand and were asked to play naturally (i.e., as during a concert performance). We used clarinet players, as their visible body movements are not directly related to sound production. The study therefore directly investigates music-related ancillary gestures.

For the experiment, we used only the first phrase of the piece (12-15 seconds), since this was the most consistent part across players and was performed without in-between breathing to eliminate this as a possible cue.

For the first condition we showed digital video recordings of the players. For the other conditions, the motion-capture data was taken to create “stick figures” of the musicians. A selection of certain data points from the players was taken to connect them with a white line building a sparse shape of the person.

The second condition presented the stick figures of the players from the same viewpoint as shown in the video recordings (90 degree from the left). In the third condition, we showed the stick figures from the front (see Figure 1). While possible side-to-side motions of the player were difficult to detect in the side view, due to the abstract and transparent presentation, in the frontal view possible front-back motions were less recognizable.



A)

B)

C)

Figure 1. Screenshots of one player: (A) video recording, (B) stick figure side view, and (C) stick figure frontal view[1]

2.1.2. Design

The experiment was designed and conducted with the Psychophysics Toolbox (PTB-3)[2] in Matlab on an Apple G4 notebook. The videos were presented on the screen with 560x420 pixel size and 25 fps. Sound was provided through headphones.

Participants had to press the space bar to start the video sequence. The movies were shown in randomized order and could only be seen once. After the sequence ended, the participants were asked to rate the *tension*, the *intensity*, the *fluency*, and the *musician’s professionalism* in a 7-point Likert scale. After the four ratings were made, the participants were able to start the next trial with the space bar. The movies were shown without repetition.

[1] The face in Figure A is blurred to hide performers’ identity, though there was no blurring during stimulus presentation in the experiment.
[2] http://www.psychtoolbox.org

The different scales of rating were introduced to the participants as follows:

- *Tension*: With this scale, a more general musical impression of the player has to be judged. The concept of musical tension is a complex phenomenon and hard to describe formally. The participants were told to rate in this scale the relaxation, stress, and immersion of the performance and the performer. A high tension refers to a feeling of excitement, whereas a low tension refers to uncertainty and relaxation. Vines et al. [10] found that different manners of movements elicit different tension ratings. Here, we wanted to find out if tension is contained in the body motion or needs other visual cues.
- *Intensity*: This scale relates to the expressiveness and the emotion. High intensity meant that the player exaggerated the performance in both fields in relation to the piece. Unexpressive and emotionless performances should be rated with low intensity. This scale relates to Davidson [5].
- *Fluency*: In this scale, participants were asked to rate the smoothness of the performance. If the player drew a clear bow over the whole phrase it should be rated with high fluency. Otherwise, a low fluency was related to a more jerky performance. Here, we wanted to see if the body motion supports the impression of fluent performances.
- *Professionalism*: Finally, the participants were asked to rate the musician's ability to play the instrument. Here, the extremes referred either to a player who seemed to be a beginner on the instrument or to a professional player. This scale measures whether high-level visual cues influence the judgment of the musicians' expertise and ability.

Additionally, the participants were told to answer intuitively. Furthermore, they were asked to use the "4" if they think the performance was played in a standard and neutral way related to the piece. The extremes of each scale were judged relative to a neutral performance by the player.

2.1.3. Participants

Ten graduate students from the McGill Music Lab attended this pilot (age 25-35). Most of them were music educated and none of them was a clarinet player. They were not paid for their participation.

2.2. Results and discussions

The mean ratings split by the different presentation styles and scales of rating are shown in Figure 2. Overall, the ratings seem to be rather similar and are distributed in a close range around the 4 as the middle answer (3.7 to 4.7). A two-way analysis of variance (ANOVA) was run for each scale for the factors presentation and player.

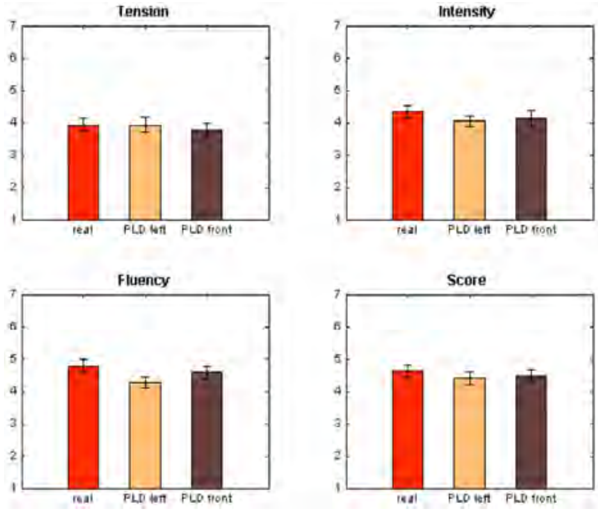


Figure 2. Mean ratings of the participants for each scale and presentation style (PLD: point-light display). Error bars depict standard error of the mean.

For all scales we found no significant effect for the presentation ($F(2,18)<3.0$, $p>0.1$), indicating that the individual ratings were not influenced by the different display styles. This suggests that the abstract display of pure body motions provides similar impressions for judging these scales to those of the real videos. Furthermore, different viewing angles did not seem to influence these experiences.

A significant effect of player was found only for tension ($F(3,27)=3.6$, $p<0.02$) and for intensity ($F(3,27)=4.6$, $p<0.01$), but not for fluency and professionalism ($F(3,27)<1.0$, $p>0.4$). This effect was mostly due to one player, who was rated lower than the others. Interestingly, for all scales there was no significant effect for the interaction of player and presentation style ($F(6,54)<0.8$, $p>0.6$), indicating that the individual player was rated in a similar way for the different presentation styles.

Although all players performed naturally, they played the piece in a very individual way. For instance, the average tempo of the performances varied considerably (Player 1: 97 bpm, Player 2: 113 bpm, Player 3: 118 bpm , Player 4: 91 bpm). Even then, the findings show that the listeners got rather neutral impressions for each player and rated them in a similar way across the scales.

To test if the different scales were judged independently, we also analyzed correlations between the scales (see Table 1). Overall, the correlation coefficient (R^2) is rather low, indicating that the scales were treated separately and did not interact with each other. For fluency and professionalism the correlation is about 0.5, suggesting that fluent and smooth performances were sometimes dependent on a higher level of instrumental expertise.

Table 1. Correlations between different scales

R^2	Intensity	Fluency	Profession
Tension	0.24	0.01	0.06
Intensity	XX	0.14	0.34
Fluency	0.14	XX	0.49

3. Conclusions

In this paper we investigated the influence of presentation style on the perception of musical performances. Participants were shown either the real video or point light displays of clarinet players. They were then asked to rate the tension, the intensity, and the fluency of the performance, along with the player's professionalism.

Overall, the results show that the different presentation styles were rated similarly, suggesting that the presentation of pure kinematic body motions of a musician seem to generate the same impressions as a real video, particularly for the scales used. Furthermore, presenting a frontal or a side view of the musician did not change the aesthetic and musical experience of the performance. This suggests that both frontal and side motions contain indicators related to similar musical impressions.

In summary, we have shown that music-related judgments could be sufficiently communicated through pure body motions of the musician. This experiment, however, still leaves open questions. Where exactly is the information located? Which motion raises what kind of experiences? How far does pure body motion effect the communication of complex aspects of musical performances? Further experiments in this project currently address these questions

4. Future work

The major goal of the project is to investigate the influence of body motions on music perception. Since this experiment has shown that the point-light displays carry sufficient information to give certain musical impressions, current studies are using these point-light displays to examine the role of different body parts and the relationship between sound and appropriate motions. Physical analyses of the body motions are also being considered in order to correlate them with perceptual ratings.

Furthermore, it is planned to use this method of musical ratings to validate the believability of virtual musicians. Mazzarino et al. [7] used quantitative gesture analysis to investigate similarities between animated characters and real recorded motions. They developed an inverse kinematics model using prioritized constraints to animate and reconstruct virtual musicians. Using this method of measuring musical impressions we aim to investigate perceptual aspects of the animation model as it is presented to listeners. Findings from this can help in designing and improving believable human computer interfaces.

Finally, this project will contribute general insights into how we perceive musical performance and whether ancillary gestures provide additional information to the listener.

5. Acknowledgments

We would like to thank Barbara Mazzarino and Ronan Boulic for helpful input and discussions. This research has been supported by the E.U. Network of Excellence on ENACTIVE Interfaces[3].

References

[1] Cadoz, C. and Wanderley, M.M. "Gesture: music", In: M.M. Wanderley and M. Battier (Eds.), Trends in gestural control of music, Paris, IRCAM/Centre Pompidou, 2000

[2] Chaffin, R. and Logan, T. "Practicing perfection: How concert soloists prepare for performance", Advances in Cognitive Psychology, 2 (2-3), 113-130, 2003

[3] Clarke, E.F. and Davidson, J.W. "The body in music as mediator between knowledge and action", In: W. Thomas (Ed.) "Composition, Performance, Reception: Studies in the Creative Process in Music", Oxford University Press, 74-92, 1998

[4] Dahl, S. and Friberg, A. "Expressiveness of musician's body movements in performances on marimba", Gesture Workshop at the 5th International Workshop on Gesture and Sign Language Based Human-Computer Interaction, Genoa, 2004

[5] Davidson, J. "Visual perception of performance manner in the movements of solo musicians", Psychology of Music, 21, 103-113, 1993

[6] Johansson, G. "Visual perception of biological motion and a model for its analysis", Perception & Psychophysics, 14, 201-211, 1973

[7] Mazzarino, B., Peinado, M., Boulic, R., Volpe, G., and Wanderley, M.M., "Improving the believability of virtual characters using quantitative gesture analysis", International Gesture Workshop, Lisbon, Portugal, 2007

[8] Scully, D.M. "Visual perception of technical execution and aesthetic quality in biological motion", Human Movement Science, 5, 185-206, 1986

[9] Thompson, W.F., Graham, P., and Russo, F.A., "Seeing music performance: visual influences on perception and experiences", Semiotica, 156 (1/4), 177-201, 2005

[10] Vines, B.W., Krumhansl, C.L., Wanderley, M.M., and Levitin, D.J. "Cross-modal interactions in the perception of musical performance", Cognition,, 101, 80-113, 2006

[11] Wanderley, M.M., Vines, B.W., Middleton, N., McKay, C., and Hatch, W. "The Musical Significance of Clarinetists' Ancillary Gestures: An Exploration of the Field", Journal of New Music Research, 34:1, 97 – 113, 2005

[3] http://www.enactivnetwork.org

Vestis: sensory spaces of mediation

Luisa Paraguai
University of Sorocaba, UNISO, Brazil
E-mail: luisa.paraguai@uniso.br

Abstract

The paper is concerned with the mobile technologies, more specifically the wearable computers, and their interventions on the users perception of body and space by demanding new behavioural codes and evoking other communication patterns. The body, nowadays mediated by those technological interfaces, has to reconfigure its own expression forms to embody hybrid dimensions of interaction. After some theoretical approaches, the vestis project will be presented as an artistic experimentation at the visible and tactile domain of those new spatial conditions.

1. Introduction

The computing technologies and networks have effectively contaminated the physical space and to understand the materiality and effect of mobile devices' use on those relations to body and space it is the main theoretical concern of this text. According to Manovich (2005) *"if the 1990s were about the virtual, it is quite possible that this decade of the 2000s will turn out to be about the physical – that is, physical space filled with electronic and visual information"*.

The technology enables users of being always connected, creating other practices of sociability by composing the urban landscape with virtual spaces. This process transforms the user relation to the space, that it is no longer physical or virtual, but hybrid. According to Hayles as cited by De Souza e Silva (2004) *"the context is becoming enfolded, so that there is no longer a homogeneous context for a given spatial area, but rather pockets of different contexts in it."* A hybrid space, thus, is not only related to the layering of digital data on physical reality. It is a conceptual space created by the merging of borders between the physical and the virtual space, due to the use of mobile technologies. (De Souza e Silva, 2004) Some data can be embedded in objects located in the space around the user, can come from the space itself, and come from global networks. *Then, the garment, the home, and the public space become sites of processing and mediation*. (Greenfield, 2006)

The technological context of wireless communication and embedded processors has been able to a transaction between the user and an information-processing system that proceeds automatically without the need for users' action, intention or even consciousness of what is taking place – what Greenfield (ibid) might call *"information processing dissolving in behaviour"*.

Several theorists have been writing about the phenomenon of wireless communication and its consequent social mobility - a fluid organization with holes not so rigid and social spaces constantly changeable. Meyrowitz (2004) proposes that the contemporaneous culture would be creating *"the global nomads"* and Urry (2000) affirms that the actual culture can be comprehended only by the concept of *"a mobile sociology"*. Then, the actual process of urban mobility despite creating body and information displacements has evoked other ways of space and time control by the juxtaposition of virtual and physical spaces. For example, someone who travels constantly, but is connected on the Web, is a mobile user, controlling and being controlled by the flux of matter and information.

2. Hybrid bodies: dimensional spaces of action

The context of mobile technology can be characterized by different levels of disconnection and connection from the actual space, running cyclically. The mobile devices do take people out of their actual context, but also bring remote people and information to their nearby context. They combine both, the centre and the periphery of users' attention and in fact provoke the constant movements of forward and back between the two conditions.

For example, thinking about mobile phones, users can cut themselves off from their environments or stay engaged with the actual context while making or receiving a phone call. The movement of in and out promotes new types of sociability and communication patterns and it transforms the senses of public and private space by creating distinct behavioural dimensions. Mobile calls can come at any time, at any place, and in the company of any number of onlookers and eavesdroppers. What is considered private conversations now can take place in public situations in which bounded intimacy are determined by characteristic bodily gestures and postures. According to Plant (2001), the receipt of a public call can be seen as a ritual behaviour which presents different levels of disconnection, characterized by three distinct situations: "*these are flight, in which users immediately move to absent themselves from their social situation; suspension, in which recipients stay put, but stop whatever they are doing for the duration of a call and effectively cut themselves off from their environment; and persistence, in which users stay put and engaged with the actual world, as far as possible carrying on with whatever they were doing before they made or took the call*".

A mobile phone user really is trying to live in two contexts, when people are frequently seen riding bikes, eating, carrying packages and talking on their mobile all at the same time. Plant (ibid) uses the word *"bi-psyche"* to conjure this sense of operating on two dimensions. It also has a physical manifestation: a whole new set of postures adopted by mobile phone users and this is the main concern of this paper. She describes the *"speakeasy"* posture - head thrown back and neck upright, *"giving out an air of self-assurance and single-minded refusal to be distracted by the outside world."* Then, there is the *"spacemaker"* posture - head bowed, with the user often walking around in a circle trying to carve out an imaginary private space. If the speaker is sitting, for example on a park bench, then this aura of withdrawal might be enhanced by a pair of feet drawn up off the ground. Or the "*spacemaker*" might turn away from the rest of the world toward a corner or a wall.

Then, it is possible to assume that body gestures have not been any more determined by local demands only; people have to negotiate the actual and remote interactions and reconfigure their body zones for this new spatial condition. For Rheingold (2005) the mobile phone use and ritual interaction are interconnected and able to create social changes and social capital.

3. Vestis

By thinking about those changes associated with mobile communication, there are several artistic experiences being elaborated, and it is possible to enumerate some that create not usual connections between people in public and urban space: Social mobiles – a collaboration project between IDEO and the artist e designer Crispin Jones, 2002 (http://www.ideo.com/case_studies/social_mobiles/menu.html), Umbrella net - collaboration between Jonah Brucker-Cohen and Katherine Moriwaki, 2004 (http://www.mee.tcd.ie/~moriwaki/umbrella/), Whisper - collaboration between Thecla Schiphorst and Susan Kozel, 2002 (http://lab.v2.nl/projects/whisper.html), Head, wearable sculpture - Laura Beloff (http://www.saunalahti.fi/~off/head/index.html).

The project "vestis" has been developed and pointed out the use of a wearable computer as a potential interface of reconfiguring the corporeal schema and lived experiences of bodily spatiality. By appropriating the technology the artist, Luisa Paraguai, intends to formalise the individual's personal boundaries aesthetically as the result of a fluid process of those body spaces contractions and expansions. Then, the embodied space can be formalised by an interactive process between the wearer wearing "vestis" and participants around him/her.

The first prototype of "vestis" has a structure with four circular nylon tubes that can expand and shrink independently; those movements give to it different shapes and contours at each moment of the dialogue between participants, now interactors, and the wearer. (Figure 1) This dynamic form is possible with a telescopic assembly of tubes that gives perimeter variations. Those variations are possible through a system of flexible screws inside of tubes, which are put in action by micro motors. A microcontroller, running proprietary software, defines and controls all these movements after getting inputs from embedded sensors that monitor participants' presence and wearer's responses.

The participants' presences are determined by their physical distances from the wearer, tracked by the back and frontal ultrasonic sensors at regular intervals of time (about 3-4 seconds). For each distance has been defined a specific clock or anti-clock rotation movement of the micro motors, which implies on a contraction or expansion independent movement of each nylon tube. The attributed values at the program can be changed, according to the artist's intention trying to evoke different approaches of the personal body zones of interactions. The closer the participants are, the smaller the "vestis" structure becomes; a "constrictive body" makes visible, materialising for instance the bodily constrictions from formal interactions. Other bigger shapes of "vestis" configure a more "expansive body", by allowing freestyle movements and body gestures patterns. Another consideration for the "vestis" shapes definition is the direct dependence of the sensor position; it means those shapes are quite different if the participants come towards the wearer or her (his) back.

Figure 1. The “vestis” structure.

Then, “vestis” formalises sensorially the fact we “have” and “are” bodies (Turner apud Low, 2003), dependent of the cultural and social relationships elaborated as spaces of negotiation. People need to wear “vestis” and experience their own body and space; then, getting involved in the work process they will perceive the sensory changes the dialogue/interaction among them can provoke. The artist assumes that reality is not a given, but reality is shaped and moulded through communication.

“Vestis" tries to play a poetic way with visual and tactile channels for experiencing “embodied spaces”, evoking participants’ engagement as an effective and affective negotiation of the use of body connections.

4. Some final considerations

More and more, the mobile devices use has been included in the people daily life, and what is observed are the stylized social behaviours being repeated frequently everywhere, which have been establishing other social bonds and personal boundaries. “Vestis” is proposed to materialize the needed reconfigurations of the bodily spatiality, by being the users, nowadays, mediated by different technological interfaces.

Another consideration for the creation of new behaviour patterns can be a tendency of modulation of technological devices – the development of pervasive networked objects. A computing and information distribution among different objects, wearable or not, can demand new functionalities, and therefore the need of new behaviours patterns elaborated by the users. According to Sterne, as cited by De Sá (2004), there is a contemporaneous way of hearing and it is articulated to the technical objects of audition as the telephone, the telegraph, the phonograph, and even the medical devices as the stethoscope. The necessities of learning process of those objects use have been transformed people’s way of hearing.

Reflecting about the embodiment of mobile devices in the daily life, it is possible to affirm that users have been performing other bodily movements, understood as dynamic inscriptions of the body. The comprehension of that body has to consider the emergence of spatialities configured by juxtaposed dimensional vectors of the body, the local and remote physical space, and the information.

References

[1] De Sá, S. “Telefones Móveis e Formas de Escuta na Contemporaneidade”, Razón y Palabra, n.41, October-November, Tecnológico de Monterrey, Campus Estado de México, 2004. [online] Retrieved on 16 January 2006 from http://www.cem.itesm.mx/dacs/publicaciones/logos/anteriores/n41/spereira.html.

[2] De Souza e Silva, A.A. From multiuser environments as (virtual) spaces to (hybrid) spaces as multiuser environments: Nomadic technology devices and hybrid communication places, The Federal University of Rio de Janeiro Press, Rio de Janeiro, 2004.

[3] Greenfield, A. Everyware: The dawning age of ubiquitous computing, New Riders Press, London, 2006.

[4] Hayles, K. How we became posthuman—virtual bodies in cybernetics, literature, and informatics. The University of Chicago Press, Chicago, 1999.

[5] Kleinrock, L. “On some principles of nomadic computing and multi-access communications”. IEEE Communications Magazine, pp.46–50, July, 2000. Retrieved on 08 April 2006 from http://www.comsoc.org/~ci/public/2000/jul/index.html.

[6] Low, S.M. “Embodied Space(s): Anthropological theories of body, space and culture”, Space and Culture, vol.6, n.1, pp.9-18, SAGE Publications, Thousand Oaks, 2003.

[7] Manovich, L. The Poetics of Augmented Space: Learning from Prada, 2005. [Online] Retrieved on 15 May 2006 from http://www.manovich.net/DOCS/augmented_space.doc.

[8] Plant, S. On the mobile. The effects of mobile telephones on social and individual life, Motorola Inc., 2001. [Online] Retrieved on 27 May 2006 from http://www.motorola.com/mot/documents/0,1028,296,00.pdf.

[9] Rheingold, H. Mobile Phones, Ritual Interaction and Social Capital, 2005. [Online] Retrieved on 01 June 2006 from http://www.thefeaturearchives.com/topic/Culture/Mobile_Phones__Ritual_Interaction_and_Social_Capital.html

Experience and Abstraction: Enaction and digital cultural practices

Simon Penny
University of California Irvine
E-mail: penny@uci.edu

Abstract

This paper is concerned with the nature of traditions of Arts practice with respect to computational practices and related value systems. At root, it concerns the relationship between the specificities of embodied materiality and aspirations to universality inherent in symbolic abstraction. This tension in embodied in the contemporary academy, as embodied arts practices interface with traditions of logical, numerical and textual abstraction in the humanities and the sciences.

The computer may be viewed as the reification of a rationalist world view in that the hardware/software binarism, and all that it entails, is little but an implementation of the Cartesian dual. Inasmuch as these technologies reify that world view, these values permeate their very fabric. Social and cultural practices, modes of production and consumption, inasmuch as they are situated and embodied, proclaim validities of specificity, situation and embodiment contrary to this order.

1. Introduction

This paper is an attempt to make explicit a set of issues regarding implicit disjunctions between technological and cultural practices. Cultural practices are traditionally often concerned with specificities of history, personality and context. They have not, in the past, been subject to evaluation on the basis of instrumental criteria such as efficiency, productivity and optimality. With the emergence of computing as a commercial and a cultural force, these values have insinuated themselves, into areas of cultural practice. Where computational technologies are engaged by social and cultural practices, there exists an implicit but fundamental theoretical crisis. An artist, engaging such technologies in the realization of a work, invites the very real possibility that the technology, like the Trojan Horse, introduces values inimical to the basic qualities for which the artist strives. The very process of engaging the technology quite possibly undermines the qualities the work strives for.

In my opinion, the full force of some of these realities is felt most clearly by the practitioner in the complex process of realisation of cultural artifacts employing these technologies. Contemporary digital arts practice is shaped, in large part, by the ramifications of the disjunctions discovered in a process where technological components formulated for instrumental ends are applied to goals which exceed these instrumental conceptions. What is called for then, is a simultaneous assessment of these values ands their implications to contexts outside their 'native' territory, and simultaneously, a reassessment with respect to these issues, of the core values, methodologies and sensibilities of the arts.

2. Cartesianism and Craft

There is, in the western academy and other aspects of western culture, a deep value which ascribes greater worth to more abstract and 'mental' work, and implicitly or explicitly denigrates work which involves manual labor and skill, and therefore devalues the people who do that work. Technical labor, crafts and trades, bodily training in sports, dance and martial arts, often require high intelligence (think of virtuoso musicianship). The rationalism of the academy is characterized by the valorization of symbolic forms of representation: textuality, logico-mathematical symbol systems, and symbolic representation more generally. The arts, at their core, bypass this translation from worldly experience of materiality to symbolic representation as alphanumeric characters. The arts is largely concerned with the way objects, forms, materials, and bodily actions can *mean*. The arts focus on immediate sensorial experience, unmediated by alphanumeric translation.

A most significant difference between computer science research and media arts practice lies in the ontological status of the artifact. As discussed above, for an artwork, the effectiveness of the immediate sensorial effect of the artifact is the primary criterion for success. It is engaging, it is communicative, it is taken to be coherent, or it is a failure. The criterion for success is performative. Most if not all effort is focused on the persuasiveness of the experience. In computer science the situation is reversed. If the physical presentation is a little rough around the edges, or even missing entire pieces, this can be overlooked with a little handwaving, because the artifact functions as a 'proof of concept' which points to the real work, which is inherently abstract and theoretical.

Computer science, as a technical discipline, reifies philosophical notions which, oddly, were already under interrogation in other disciplines prior to its formation. The conception of intelligence in computational discourses is rooted in an early-mid C2Oth approach valorizing mathematico-symbolic problem solving. This issue is of great significance in the current discussion, as the kinds of intelligences which enable the arts and cultural practices are among those exclude from the mathematico-symbolic conception. Handwork can involve high intelligence and sensibility. But that kind of intelligence – embodied, kinesthetic and multi-modally sensorial intelligence, tends to be irreconcilable with textual, alphanumeric logico-symbolic forms of work. Contrarily, the process of translation from the abstract to the concrete is an exercise of high intelligence, and valuable knowledge and insight is drawn from actual manipulation of matter, as opposed to talking about it or using pre-constructed simulations.

3. Embodied and situated practices

It is odd that we are prepared to accept as generally useful, a machine system which is only capable of interpreting as input, linear strings of alphanumeric characters. The machine knows nothing of the world, except that which a human predigests and feeds to the machine as alphanumeric strings. Such a system is excellent for doing arithmetic and accountancy, calculating tide and firing tables, storing and retrieving textual records because these practices have already been abstracted into formal mathematico-logical representations and organizational and cataloging systems generations before the machine existed. Indeed, the machine is well attuned to these practices because the formalisms upon which the machine is based, and the formalisation of those organizational practices arise from a common root.

Reflect on the larger historical arc, beginning, as AI practitioners like to do, with Descartes and the establishment of rationalism. Here, in broad terms, we see the success of attempts to categorise and organize the world according to mathematico-logical ordering systems. Subsequently, we see the development and increasing sophistication and elaboration of mathematically based techniques for designing and building machines: engineering and the industrial revolution. This paradigm gains momentum as electricity, radio, telegraph and related technologies arise, and coalesce as electrical engineering and electronics, during a time when engineering itself is being reconfigured as an increasingly analytic and mathematical discipline. From this technical base arises electronic, and digital computing. The implementation of Boolean logic as electronic machine was the foundation upon which programs like logic theorist ran. So it should be no surprise them that such technology was found to be highly amenable to the automation of mathematical logic, and by the same token, it explains why problems outside that realm have been found so intractable.

Our world is replete with complex cultural and social practices in which the calculation, storage and retrieval of data play a vanishingly small part, and in which spatial awareness, texture, gaze, gesture, tone of voice, perceptual integration, active sensing, kinesthetics and proprioception (all sensibilities outside the ken of the computer) play key roles. What this means, in effect, is that the technology to which we are encouraged to apply to these functions is incapable of sensing or measuring these qualities (I hesitate to even call them variables). In effect, the conventional PC is a filter which filters out all aspects of our complex embodied intelligence except that small part which can be encoded as strings of alphanumeric characters.

Rhetorics of computing, both marketing rhetorics and the more complex and subtle characterisations of the computational in literature and film, commonly contain extropian and anti-corporeal sentiments which imply that human experiences which are not amenable to serial Boolean logical expression are somehow irrelevant. Surely this should be an issue of greatest concern to practitioners and theorists of embodied practices, yet there is an almost entire absence of informed critical assessment of the relevance of such a technological paradigm to activities like, for instance, choreography, painting, cooking, sailing, clinical diagnosis or physical therapy.

The simple fact is that media arts employ technologies designed for instrumental purposes – automation, accountancy, archiving. It cannot be asserted that artistic needs and purposes were ever considered in the design of the basic technologies. It follows then that existing computer technologies are unlikely to be optimally appropriate for such applications. This is unlike, for example, the evolution of the medium of oil paint, which was developed over generations specifically for the task of painting pictures. A machine designed for manipulating strings of alphanumeric characters may simply not be relevant to certain human tasks – why should we assume it should be? Why should we be at such pains to deny the obvious fact that our intelligence and our embodiment are precisely attuned to each other, through childhood development as well as through evolutionary process? Our intelligence is expressed in all modes and all combinations of modes of our lived physical being. Yet we are increasingly naturalized to the idea that we should be ready to translate any sort of human notion or practice, into keystrokes, in order to make in acceptable to this device. Not only is it absurd that such an expectation be attached to such a purportedly marvelous technology, but it relegates any human quality not amenable to such processing to oblivion or irrelevance.

4. Information and meaning

Fundamental to CS is the idea of information, and the idea that information exists, or can exist, in some abstract non-material realm, separate from and independent of, its material substrate. This is an (inherently Cartesian) assertion and not a self evident truth. As a structuring assumption it is ripe for critique. As such, it has permitted the sorts of advances compatible with the paradigm, but, equally, has excluded entire avenues of research. Due to the elaboration of this paradigm, an ontological drift in the term 'information' has occurred over the past half-century under the influence of the development of techniques which utilize Boolean operations in a von Neumann architectures. I suggest that the discipline is structured by an informal working definition which is not unproblematic because it confuses 'information' with 'computability'. 'Information' has been formalized as quantifiable and logically manipulable (Shannon), and hence, *information* which is not quantifiable and logically manipulable is no longer information. We must examine the value structure thus created, because if logical manipulability is valorized, then vast realms of human practices are hence devalorised.

5. Objectivity

One of the large trends in western thought over the last century, felt equally in the sciences, in the humanities and the arts, has been the challenges to the presumed authority, validity or even possibility of objective knowledge or a detached objective viewpoint. As Heinz von Foerster remarked "Objectivity is a subject's delusion that observing can be done without him." The culture around computer science, like any other academic discipline, has its inconsistencies and oddities. These include subscription to an unreconstructed Cartesianism and unreconstructed Objectivism, explicit in the 'gods eye view' often encountered in software and systems. Enactive and situated theories of cognition and phenomenological critique of AI (Dreyfus, Suchman, Varela, Lakoff and Johnson, et al) exposed a platonic and top down spirit in that enterprise and the school of cognitive science associated with it, and led to a recognition of the relevance of theories of situated and embodied cognition. This opened a way for more subjective and less autocratic modes of technical practice (Brooks, Maes, Agre, Horswill and Chapman et al). David Marr begins his well-known 1982 book on vision with the statement that "vision is the process of discovering from images what is present in the world, and where it is". But in human and animal biology, the study of perception as a one way process, an of methods in which are clinically isolated from lived experience has given way to the conceptualization of *active sensing*, which asserts the importance of examining the kinesthetically engaged, temporal coupling of sensing and action.

6. Generality and Specificity

A fundamental commitment of computer science is that of the *General Purpose Machine*. From the outset, generality was taken to be desirable, for reasons which are unassailable in formal terms. The unquestioned axiomatic acceptance of the concept of generality as being a virtue in computational practice demands interrogation, especially when that axiomatic assumption is unquestioningly applied in realms where it may not be relevant. Indeed, the fact that the idea of the universal relevance and validity of the concept of generality is rarely asked; itself suggests fertile ground for interrogation.

Historically one can identify a two-stage process of elision and reification, related to the economic principles of the computer industry and the rapid uptake of the computer in diverse socio-cultural contexts far from the original applications of the machine. The first stage was the transfer of the notion of 'general purpose' to the beige colored box and its big vacuum tube appendage. The idea of generality, entirely substantiable in formal mathematical terms, became thus attached to a physical commodity. The notion of generality thus offered justification for highly profitable strategies of consumer commodity. The casualties of this capitalist sortie are seldom discussed: if all uses for the computer could, theoretically, be contained by alpha-numeric desk-work, the other sorts of human practices which were not compatible with that particular work culture, or not identified as profitable enough sectors to justify trhe investment in software tool development; had to reshape themselves or suffer the stigma of remaining uncomputerised.

The world was thus divided into the computerized and non-computerised realms, and caché and advantages flowed to the computerized practices, in popular culture, which was itself increasingly defined by and located in digital practices; as in the academic and research worlds, where computerized or computerizable disciplines were able to access comparatively huge funds (much of which flowed directly back to the computer hardware and software industry). The result of this trend was that all sorts of human practices for which the computer, as formulated by the industry, was not ideally conformed, often then bent and reconfigured themselves to adapt, often at a significant cost to the integrity of the practice and its sensibilities and knowledge base. This process is observable in diverse fields and disciplines over the last quarter of the C20th, from engineering to the arts, but it is in the arts that such trends are particularly stark. This is because, as argued, the arts the practices rest on such profoundly different foundations, both historically and theoretically. This then is the core of my argument: artworks are made by individuals of particular physical conformations, with particular perceptual and physical skills, immersed in specific cultural and historical contexts.

7. Embodiment, situation and tools

Tools are specific to functions. There is no such thing as a general purpose tool. Every craft has a range of specialised tools. The skilled craftsman is highly discerning about matching a task to a tool. The notion that generality is a virtue is opposed to a generally accepted notion that there is a tool for every job and a job for every tool. Contrarily, informed by the dual evil motivations of user-friendliness and generality – software tools seek to reduce the diversity and specificity of individual and cultural motivations and world-views: user friendly software tools make easy (generalisable) tasks easier and difficult (more specific) tasks more difficult. In opposition to the ideology of generality, one might propose that art is naturally Pataphysical. An artwork is deemed to be excellent if it addresses a particular situation with persuasive precision. That is, by a subtle combination of the signifying potential of spatial organisation, materials, sounds, images and user dynamics; a coherent experience is generated which leads the audience/user into a particular realm of interpretation. An artwork is successful to the extent that it is specific. Generality is not a virtue in the Arts. Generality and affective power seem to be mutually exclusive. It's hard to imagine what a general purpose artwork would be like, unless it was one of those generic and vacuous hotel room pictures, whose work is to proclaim a respect for art on behalf of their owners, while safely avoiding the danger inherent in actually being art. This is the fatuous conundrum at the root of the myriad of techno-cultural projects which attend to and intend to automatically generate cultural artifacts. The notion of the general purpose machine has indisputable power and relevance in its place. But we must be wary of the drift of axiomatic assumptions which can flow from a paradigmatic technology of both rhetorical and economic power.

8. Abstraction and Simulation

The tension between the power afforded by abstraction, and the simultaneous loss of precision, is explicit in the case of simulation. By the same (fix it in software) reasoning, computer simulation of real world contexts must be regarded with some reservation. Any simulation tool is itself a design artifact, and depends for its representational accuracy on several factors. First, that the designer correctly identified all the relevant physical effects. Second, that such physical effects are amenable to algorithmic representation. Thirdly that these representations are accurate and of adequate resolution. Forth, that all possible interactions of these relevant factors were appropriately calculated and represented. Certain kinds of physical phenomena, particularly those manufactured to reliably embody and express a mathematically simple physical process are more simple to simulate. The behavior of a tree in a storm, or the turbulence of water on a ships hull demand more complex computation, or may be inherently uncomputatable. Here the isomorphic loop of industrialism and engineering stands out in stark relief. A gear train or resistor-capacitor network is easily simulated because these things are themselves produced to embody behavior easily described in newtonian terms. As Hamlet noted, *There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy*.

Conclusion

If the thesis of this paper is taken to be valid, at least in part, then several paths of action are called for. The first is a thoroughgoing assessment of the effects of the computational paradigms on cultural practices, this is both a theoretical inquiry and a context for historical work and case studies. Practitioners are duty-bound to assess the values inherent in the technological tools they employ, lest they sabotage their enterprise. If and when these ontological booby traps are identified, a new mode of technology development is called for: the imagining, design and development of tools consistent with the values which underly and shore up the practice itself; always allowing for the possibility that any form of technology *could be* antithetical to, or destructive of, the cultural enterprise. The other, alternative and parallel path is the negotiation of new cultural practices native to the new technologies, in a process which intelligently and attentively assesses the potential disharmonies between the artistic goals and the qualities of the technologies.

Both these kinds of design processes must address the technologies at a plethora of different levels, from the smallest component level to entire devices, from implicit entailments of programming languages to dynamics of the interface and interaction, and everything between.

Notions of Quality; factors influencing why artists choose their media.

David Prytherch[*] Christos Giachritsis[*†]
(*)User-lab, Institute of Art and Design, Birmingham City University, UK
(†)SyMON Lab, School of Psychology, Birmingham University, UK*
E-mail:david.prytherch@bcu.ac.uk, c.giachritsis@bham.ac.uk

Abstract

This pilot study investigated human sensitivity to vibration arising from tool interaction with a material, specifically artists' perception of different types of pencil lead. We have shown that expert human observers were able to tactually discriminate between different lead types used for drawing. They found it easier to discriminate H-range than B-range leads against a standard HB. Tribological tests on the leads used in the experiment show that 2H and 4H leads generate approximately 40% and 70%,, more friction than an HB, respectively, while B-range leads generate very similar friction as the HB. This may suggest that participants responded to frequency differences generated by different frictions and H-type leads may produce higher frequency stimulation to which the human tactile system is very sensitive. To compare our results with those obtained in previous studies we intend to establish the frequencies generated by the leads used in the present study (in progress).

1. Introduction.

Artists and contemporary craft 'makers' are frequently very specific in their choice of medium, (the material they work with) and the tool or tools with which they manipulate this material, for both are of critical importance in defining the quality and aesthetic of the outcome. Whilst some factors influencing their choices may be relatively easily described in terms of visual aesthetic or practicality, many others remain tacit and are rarely discussed and it is suggested that the basis for these lies within the domain of haptic perception.

These general principles of choice apply throughout the range of fine and applied arts. A graphic designer may have a preference for a particular paper and marker, pencil, charcoal, ink etc. [1], a blacksmith may prefer a particular kind of steel and hammer with which to beat it. [2] These are judgements of "quality" and the decisions made at the outset of the process have a profound effect both on the perceived quality of outcome and the progress of personal skill development. Much of this haptic engagement with the process of making art seems to occur at a preconscious level, the level of tacit knowledge, and is symptomatic of why it is so rarely directly discussed. This engagement may be so subtle that even when artists are actively reflecting on the sensory and perceptual constituents of their processes, they may still remain largely unaware of the pervasive role that their haptic sensibilities are playing. [3]

Many sensory factors relating to material intimacy seem to be held in common amongst makers, no matter what the material of choice or the types of processes involved. Attributes such as hardness, softness, smoothness, texture or weight may be expressed by makers as contributing to the allure of their area of expertise. It has been this author's experience that detailed and intimate physical criteria are rarely discussed by artists and makers and seem little researched by the scientific community. However, amongst makers, there seems a tacit acceptance and understanding of the significance of such considerations though these characteristics may seem commonplace and unimportant to others.

Figure 1. Stage 1 – La La La. A dancing sculpture in 25mm float glass and 24% lead crystal by David Prytherch.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

A glass maker will express a preference for a particular type of glass, sometimes going so far as to make it “from batch”, that is, from a precise formulation of specific ingredients in order to produce a “metal” that has specific visual and handling properties. For the engraving and carving of glass, a full or half lead crystal tends to be preferred since it has greater optical clarity and is softer to work, though brittle, and these physical characteristics allow greater control of the process. This in turn produces a satisfaction in the process which is considered fundamental in the development of dextrous skill and motivation, experiences leading directly to material/tool “embodiment”. [4]

A painter may express a preference for a particular combination of surface material and paint, canvas and oil paint or board and acrylic etc. Some of these choices are practical, some aesthetic, but the experience of practice suggests that there is a significant tacit component within this choice, often simply expressed as "it just feels right". This tacit component can be so powerful that it will take precedence over other, perhaps more practical considerations such as discomfort or unpleasantness associated with certain aspects of the process.

Figure 2. "Midnight in the Garden of Science and Industry" – Alessandra Kelley
Egg Tempera on Claybord - 18"x24".

The American artist Alessandra Kelley specialises in painting with egg tempera. This is a notoriously painstaking technique in which the artist must first prepare a stiff board, such as wood or clay first by sizing, applying a thin layer of rabbit skin glue which acts as an absorption barrier. Then the board must be primed with chalk gesso to provide a suitable base on which to apply the paint. The artist then prepares the egg medium which must be fresh since it will only keep for up to 3 days in a cool environment. This is made from egg yolk from which every trace of egg white and membrane has been removed, mixed with a little water. Finally the pigments themselves are prepared by grinding on a glass plate and mixing with a little water and the prepared egg medium. The painting technique itself is also unusual since the paints are quickly absorbed into the gesso background leaving no time to move them about or to blend visual effects in the same way as you could, for instance, with oil paints.

Nevertheless, there is an intense satisfaction intrinsic in the process which is quite separate from the equally important issues of communication. Kelley states,

“I like the feel and the satisfaction of laying down the paint that I have made myself, to bring an image or idea to life. I have spent a long time coming to understand my materials, and I enjoy being able to manipulate them to produce effects. At the same time, I also appreciate the unexpected results that can show up (the astonishing variety of pigments, for example, which can be used to produce bright violet). I also like to play with paint application. Obviously, egg tempera has some physical limitations (chiefly involving paint film thickness), but within those limits is much room for experimentation. I've applied paint with sponges, toothbrushes, paper, beautiful Russian sable brushes, nasty old bristle brushes, and (gloved) fingers. I've scraped, sanded, scratched, and stencilled. I've learned to be fearless about technique towards the final effect. [5]

The influence of aspects of haptic perception is recognisably widespread in arts practice and is acknowledged by both practitioners and teachers as having a direct relationship with the quality of the creative outcome. “The most beautiful and lively drawing will be the one in which the hand and eye are most perfectly coordinated with the sense of touch: this coordination can only be achieved with long discipline and practice.” [6] Scientific studies of this fundamental and formative influence on creativity have been few, though some detailed studies of how people draw, the nature of their pencil strokes, (starting location, direction of movement etc.) have been conducted, notably the work of Van Sommers, [7] yet formal research on the perceptual basis for media preference decisions is sparse.

Formal research on the perceptual basis for these preference decisions is sparse and it seems generally assumed that the criteria for preference is predominantly visual. A general hypothesis is proposed that though the initial judgement of “quality” may be visual, the haptic properties of the media/tool combination are of great significance in the decision making process, and may in some instances possibly be dominant. The term “quality” in this instance is taken to include both notions of material properties or characteristics and the associated “value” judgement of certain properties being “better” than others. This latter notion presupposes an experienced awareness of both the intended process and desired outcome and it can be stated that any judgements based on criteria of “fit for purpose” are predicated on this type of developed tacit knowledge.

2. The Pilot Study.

In order to investigate this, we have initially focussed on the most basic of artistic processes, drawing at its most fundamental level, simply making marks on paper. We conducted a pilot study to investigate human sensitivity to vibrations caused by manipulating a tool under natural testing conditions. For the purposes of this study, we tested sensitivity of artists to different types of pencil lead on a standard grade of 80gsm copy paper which has a medium to fine texture. As part of their studies, Fine Art students train in drawing using a variety of drawing tools, including pencil leads. Most of the students develop a preference for a specific pencil type depending not only on the visual end-product but also on the way the pencil handles and feels during drawing. For example, hard (H-range) pencils produce fine, light gray lines and feel ‘harder’ while black (B-range) pencils produce bolder, darker lines and feel ‘softer’. The tactile qualities of the pencils come from the different quantities of clay and graphite they contain. Graphite is a lubricant generating minimum friction while clay is a very hard material generating maximum friction. Historically, there are no industry standards for these grades for mixing graphite and clay to produce different types of lead, though those that are marketed are first tested manually by expert quality controllers. B-range leads contain more graphite than clay while H-range leads contain more clay than graphite. If Fine Art students chose their pencils partly on the basis of how they feel, then they should be able to detect the different tactile signals produced by the different types of lead; that is, they should be able to judge whether a pencil is ‘harder’ or ‘softer’ based on tactile input alone.

Five different STAEDTLER® Mars® carbon lead types were used covering an essential range of soft (B-range) and hard (H-range) leads: 5B, 4B, 2B, 2H and 4H. Each lead was placed in a STAEDTLER® Mars® technico 788C leadholder for handling. The type of lead (i.e., No + range) depends on the proportion of graphite and clay used to make the lead. Paper scientists at the University of Manchester carried out tribological tests on the leads used in the study to find the friction they generate.

2.1. Psychophysical procedures.

A 2AFC paradigm was used to measure sensitivity to different types of pencil. Each type was tested against a standard HB type. The subjects were instructed to draw a scribble consisting of forward and backward hand movements (Figure 3) on a normal A4 printer paper (80g/m^2). The presentation sequence of test and standard leads was altered from trial to trial.

Figure 3. Typical scribble drawn by the participants during testing.

2.2 Experimental Sample.

Five undergraduate students of the Faculty of Art and Design, Fine Art department of Birmingham City University aged between 19 and 21 participated in the experiment.

2.3 Results.

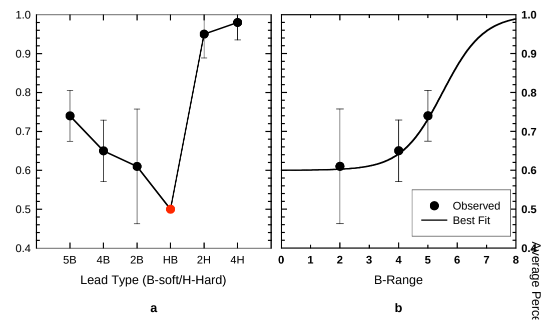


Figure 4. Overall performance across all lead types tested (a) and best fit for the B-range (b). Error bars represent 95% confidence limits.

Participants were able to discriminate between HB and any H-range pencil that was used in the study. However, discrimination was more difficult when the subjects compared HB with B-range pencils. Figure 4a shows the average percentage correct responses across all lead types tested. Figure 4b shows the psychometric function for the B-range suggesting that reliable discrimination (75% correct responses) would be achieved with a 6B lead. Unfortunately, this type of lead is not produced anymore and therefore could not be tested.

Tribological tests, which were carried out on the lead range used in the present study, revealed that lead resistance increases exponentially with lead hardness, which was expressed as following,

$$Hardness = \frac{Clay}{Graphite} * WaxWeight$$
,

3. Conclusions and discussion.

The study has shown that expert human observers are able to discriminate between different types of lead used for drawing. They found easier to discriminate an H-range than a B-range lead against a standard HB. A tribological test on the leads used in the experiment has shown that H-range generates more friction than the B-range. In fact, 2H and 4H leads generate approximately 40% and 70%, respectively, more friction than an HB, while B-range leads generate very similar friction with the HB lead. This may suggest that participants responded to frequency differences generated by different frictions. Friction generated by H-type leads could produce higher frequency stimulation to which human tactile system is very sensitive as previous studies have shown [8] [9] However, in order to compare directly our results with those obtained in previous studies it is necessary to establish the frequencies generated by the leads used in the present study (in progress).

3.1 Further Studies.

This pilot study provides us with a baseline for further experiment and establishes that our perceptual systems are indeed capable of recognising remarkably small differences in stimuli. Drawing (mark-making) is an interesting example of sensory collaboration and we naturally and automatically ascribe a basic judgement of quality to its outcome. We know whether we like it. It is a highly skilled activity, yet we can all do it and develop our skills from a very early age. Vision seems to be dominant, but it is likely that, like many other human activities, it is a product of a complex multi-modal perception, in this case probably predominantly visuo-haptic but possibly including auditory elements. An initial analysis would suggest that vision defines the environment where haptics will operate in this, and like other skilled processes, provides feedback and confirmation of process progress [4]. Yet the process itself may be predominantly haptic and begins with the artists' choices of material and marker. The choices made at the outset govern the nature and appearance, what an artist would term "the aesthetic", of the outcome. Quality of outcome in this instance seems likely to be judged by visual characteristics yet how many of these visual characteristics may also be haptic in nature. We often speak of "textures" and "grain" and other qualities relating to the "feel" of a material or object when we are describing its visual appearance.

A structured series of experiments, interviews, observations and analyses are planned to investigate not only the perceptual factors that contribute to artists' choice of media but the contribution these choices may make to their perception of the outcomes and the perceptions of subsequent viewers of the artworks. The study will integrate experimental, observational and analytical procedures drawn from a range of disciplines spanning arts, humanities and sciences. It is anticipated that this system of analysis and synthesis will yield much information of general and specific value in the study of human perception, visual and haptic, together with insights into the relationship between the artist, his/her tools and materials and the creative process. The research will lead to a greater understanding of factors influencing successful design and will provide some insight into how decorative and applied art practice may differ from fine art.

The results of these experiments will be of value not only to researchers concerned with fundamental human factors, but also to the art and design community, manufacturers of papers, markers, natural media digital input devices, and consumer objects where perceptions of "quality" are considered significant.

References.

[1] Akyuz, U. Creation of a Model for 2-Dimensional Pencil Marks. PhD Dissertation, De Montford University, 1995.

[2] Prytherch, D. Haptic Feedback in Art Making Processes with Particular Reference to Computer Interface Design, PhD Dissertation, University of Central England, 2003.

[3] Prytherch, D. & Jerrard, R., Haptics, the Secret Senses; the covert nature of the haptic senses in creative tacit skills. *Proceedings of Eurohaptics 2003*, 384-396, 2003.

[4] Prytherch, D. & Jerrard, R.. The Visualisation and Making of Sculpture and its Potential Implications for Computer Interfaces and Three Dimensional Modelling, In Baber,C., Faint, M., Wall, S., & Wing, A.M. (Eds.), *Proceedings of Eurohaptics 2001,* 135 -137, 2001.

[5] Kelley, A. *Personal communication.* 2006.

[6] Laning, E. The Act of Drawing. McGraw-Hill, New York. p.16, 1971.

[7] Van Sommers, P. *Drawing and Cognition*. Cambridge: Cambridge University Press, 1984.

[8] Brisben, A.J., Hsiao, S.S., and Johnson, K.O. Detection of vibration transmitted through an object grasped in the hand. *Journal of Neurophysiol*ogy, 81, 1548-1558, 1999.

[9] Yang, S., Tan, Z. H., Buttolo, P., and Johnston, M. Detection of Torque Vibrations Transmitted Through a Passively-Held Rotary Switch. *Proceedings of EuroHaptics 2004*, 217-222, 2004.

Why do we need 5-DOF force feedback? The case of violin bowing.

Erwin Schoonderwaldt Stephen Sinclair Marcelo M. Wanderley

Input Devices and Music Interaction Laboratory, McGill University
Centre for Interdisciplinary Research in Music Media and Technology

E-mail: {schoondw,sinclair}@music.mcgill.ca, marcelo.wanderley@mcgill.ca

Abstract

The analysis of gestures associated with playing musical instruments can provide insight into the haptic cues used to exercise control over acoustic parameters. In this study, forces acting on the bow as experienced by the violinist are described based on an analysis of bowing gesture motion capture data in crescendo-decrescendo notes, finally leading to a formulation of criteria for haptic virtual bow controllers.

1 Introduction

In playing a musical instrument, the player exerts control over a set of sound-related parameters which determine the nature of mechanical vibrations induced in some physical body. It has been shown that the bowing gesture in particular allows very accurate control over many timbral aspects of the resulting sound wave [4, 7]. Cadoz [2] observed that a sound is defined by the producing musical gesture, and that this gesture is accompanied by inherent haptic feedback, which is what allows the player to exert such fine control. However, in a bowed instrument, the control parameters (human-bow interaction) are once removed from the sound parameters (bow-string interaction), and so it is interesting to look at how forces are transferred indirectly from the sound producing mechanism through the bow to the player and vice-versa. An examination of the movements of the bow during violin playing have allowed us to identify correlations between sound and control parameters, while further analysis has revealed several directions of haptic forces at work throughout the gesture.

O'Modhrain [6] has used a 2-degree-of-freedom (DOF) haptic device to experiment with the effects of friction on bowing, and found that, although a Helmholtz friction model had little effect on bowing performance, players preferred the presence of friction over the non-friction case. This could be interpreted to imply that the force feedback of the virtual wall used for bowing may have provided adequate information to the player for execution of the bowing gesture, while the precise nature of bowing friction, much of which is filtered out through the bow, plays a less important role than the net resistance experienced by the player's right hand.

We will discuss some additional haptic cues, derived from the previously-mentioned control parameters, which provide DC-frequency force feedback to the player. These cues exist both in linear directions as well as in torsional forces, felt by the player's bowing hand. Some of these cues, such as detents, have previously been explored [5]; however, we will discuss several additional force vectors observed, illustrated by an analysis of bowing gestures during crescendo-decrescendo bow strokes. Finally, we will discuss the implication on hardware requirements for accurate haptic rendering of the bowing gesture.

2 Bow coordination in violin playing

2.1 Measurement of bowing gestures

The position and the orientation of the bow and the violin were measured using a Vicon 460 optical motion capture system. Six cameras (type M2) were used, four in front of the player and two behind. During the measurements five markers were placed on both the violin and the bow. In order to measure the full orientation of the bow, two of the markers were placed on small antennas mounted on the stick, to avoid collinear placement of the markers. During the calibration extra markers were placed to indicate the positions of the bow-hair onset at the frog and the tip (bow) and the positions where the string crosses the bridge and the nut (violin). During the post-processing these positions could be reconstructed as virtual markers for all recorded trials.

The local reference frames of the violin and the bow are shown in Fig. 1. The respective origins correspond to the virtual markers on the bridge (violin) and the frog (bow). The axes were chosen to represent physically meaningful directions, e.g. along the string (y-axis of the violin) and the bow hair (x-axis of the bow).

Bow velocity was calculated from the relative velocity of the bow and the violin, projected on the local x-axis of the bow. The bow-bridge distance was calculated as the distance between the bow-string contact point and the local origin of the violin.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22th, 2007

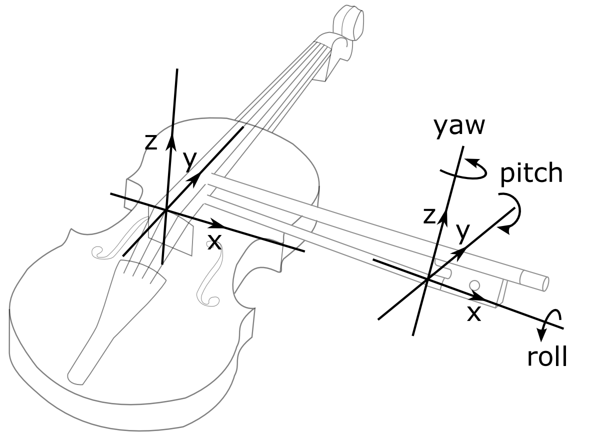


Figure 1: Local reference frames of the violin and the bow. The bow angles skewness (yaw), pitch and tilt ($-$roll) correspond to the orientation of the bow relative to the violin.

The bow angles (skewness, pitch and tilt) were calculated from the orientation of the bow relative to the violin. In the chosen geometrical representation of the violin and the bow, these angles have a direct relationship with the Tait-Bryan (or Euler) angles, obtained by respective rotations around the z-axis (yaw), y-axis (pitch) and x-axis (roll) in the local frame of reference (in this case the bow, with the orientation of the violin as starting point). The bow angles are then defined as follows: skewness (angular deviation of the bowing direction) equals yaw, pitch (angle associated with string crossings) equals pitch, and tilt (tipping of the bow, so that the bow hair is no longer flat on the string) is defined as negative roll. According to the latter definition tilt is positive when the stick is rotated away from the player, which is normal in violin playing.

Bow force was measured using a custom made sensor mounted at the frog of the bow. The sensor consisted of a spring leaf equipped with strain gages on both sides. The sensor was mounted so that the spring leaf was pressed against the bow hair, measuring the flex of the bow hair. The sensors' output signal was calibrated by pressing the bow on a calibrated force transducer at different bow positions. A calibration function was fitted to these measurements, yielding the coefficients needed to calculate the bow force as a function of bow position.

Mocap data was recorded at a frame rate of 250 Hz. The force sensor signal was measured at a sample rate of 10 kHz on a different computer. The measurements were synchronized by generating a pulse at the start of each trial using a pushbutton. The pulse was recorded in separate channels on both devices, enabling a posteriori alignment of the data.

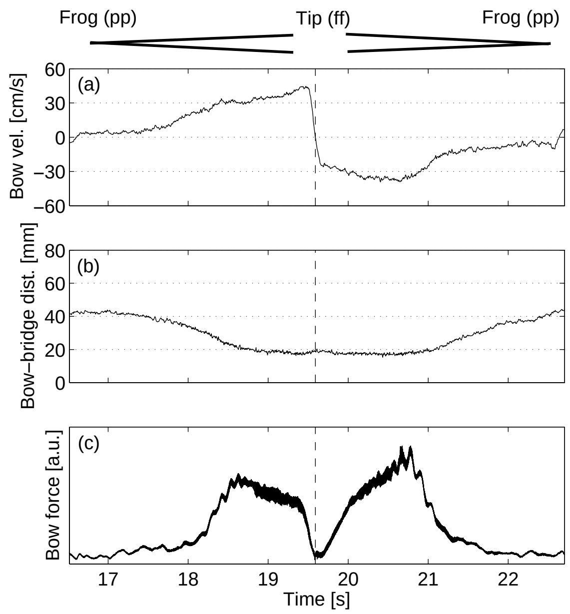


Figure 2: Coordinated change of the main bowing parameters during crescendo-decrescendo tones (down-up): (a) bow velocity, (b) bow-bridge distance, and (c) bow force. Bow position and dynamic level is indicated above the panels. The change in bowing direction (from down to up) is indicated by the dashed vertical line.

2.2 Crescendo and decrescendo notes

Crescendo and decrescendo notes form a good example of the coordination of bowing parameters [1]. An increase of the perceived loudness of a tone mainly involves two factors: (1) an increase of the amplitude of string vibration, and (2) reinforcement of the higher harmonics of the tone (timbre). In terms of bowing parameters the first is achieved by increasing bow velocity and/or decreasing bow-bridge distance (the amplitude of the string vibration is proportional to v_B/β), while the second is mainly achieved by increasing bow force.

In practice, a change of dynamic is mostly a combined effect of changes of these bowing parameters. This is clearly illustrated in the following example. A violinist (advanced music student) was asked to play a series of long crescendo-decrescendo notes, first down-up (crescendo during down bow) and then up-down (crescendo during up bow). All notes were played using the entire length of the bow. In Fig. 2 the bow velocity (a), bow-bridge distance (b) and bow force (c) are shown for one down and up bow somewhere in the middle of the series of notes. It can be seen that the bow velocity was low in the beginning of the down bow (pp) and increased until the change of bowing direction (ff). The up bow started with a high bow velocity (ff) and decreased towards the end of the note (pp). Simultaneous with the increase of velocity during the down bow, the bow was

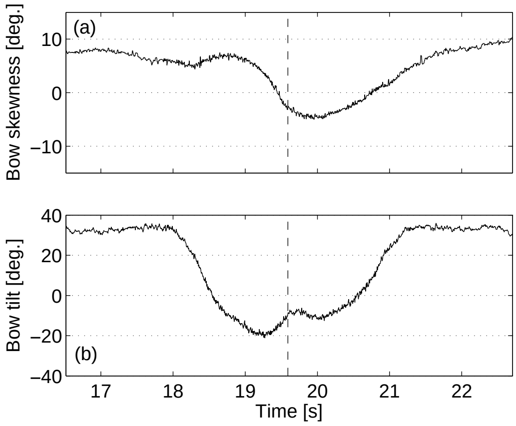


Figure 3: Bow angles during crescendo-decrescendo (down-up): (a) skewness, and (b) bow tilt. The change in bowing direction is indicated by the dashed vertical line.

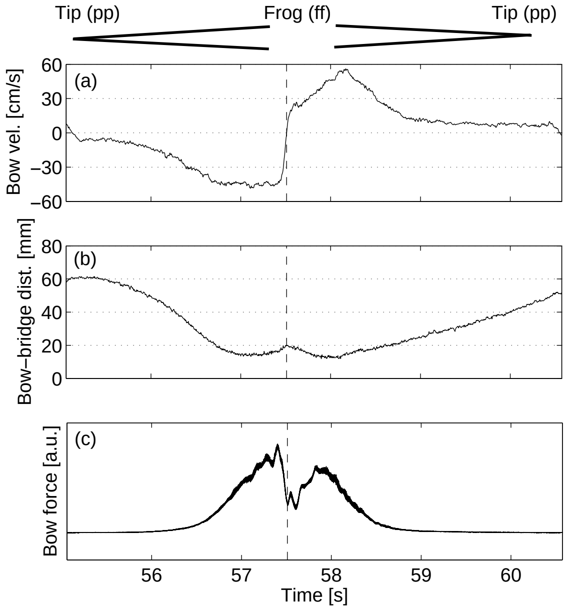


Figure 4: Coordinated change of main bowing parameters during crescendo-decrescendo tones (up-down): (a) bow velocity, (b) bow-bridge distance, and (c) bow force.

moved closer to the bridge and the bow force was increased. During the up bow (decrescendo) the bowing parameters changed in the opposite direction.

Figure 3 shows bow skewness and tilt during the same bow stroke. Arguably, tilt has less influence on the sound than the above-mentioned main bowing parameters (see e.g. Schoonderwaldt et al. [8]) and skewness has acoustically at most a deteriorating effect on the sound. However, as *control* parameters they can play a considerable role in violin playing. In Fig. 3 (a) it can be seen that the skewness was mostly positive during both the down and the up bow, i.e., the bow was angled away from the player. This causes a drift of the contact point, due to the stick-slip interaction between the bow and the string, facilitating the (continuous) change of the bow-bridge distance.

Also, the bow tilt (Fig. 3 (b)) was changed during the bow stroke, apparently in coordination with bow force. The relation can be explained as follows: in order to play softly close to the frog, tilting the bow facilitates the control of bow force as a smaller portion of the bow hair is in contact with the string. As a result the bow becomes more compliant reducing the effect of small bowing inconsistencies. However, at the tip, a high bow force is more easily obtained with the hair flat on the string.

Figure 4 shows the crescendo-decrescendo in opposite bowing direction (up-down). Comparison with Fig. 2 reveals that a similar strategy was used to achieve the changes in dynamic: the bowing parameters show the same trend except that the bow velocity is opposite in sign. In contrast, the control parameters shown in Fig. 5 were markedly different than the control parameters in the opposite bowing direction (Fig. 3). The skewness (Fig. 5 (a)) was negative (angled towards the player) for the same reason as before: the direction of the drift of the contact point changes with bowing direction. This is a strong indication that skewness—in these particular cases—was used as a control parameter to facilitate a continuous change of bow-bridge distance.

A different behaviour of bow tilt (Fig. 5 (b)) was also observed: in the second example, the tilt angle was overall much smaller. Only close to the frog the bow was slightly tilted, in this case probably to facilitate the bow change. The relation with bow force is less important here, as it is much easier to exert a high bow force close to the frog.

3 Forces acting on the bow as felt by the player

The forces exerted on the player's bowing hand during crescendo can be decomposed into Cartesian force vectors as seen in Figure 1. As described in section 2.1, the bow may be oriented in space according to three axes: pitch, yaw, and roll. Each of these orientations may be associated with a torque felt by, and exerted by, the player.

During a bow stroke, there is a non-zero net displacement of the string in the direction of bowing (x-axis), due to static and dynamic friction. This friction is modulated by the bow pressure, which is a downward force (z-axis). This is complemented by the string's upward normal and spring constant which counters the force of gravity and reacts to the bow pressure.

However, since the bow's frog lies some distance from the bow-string contact point, the downward bow

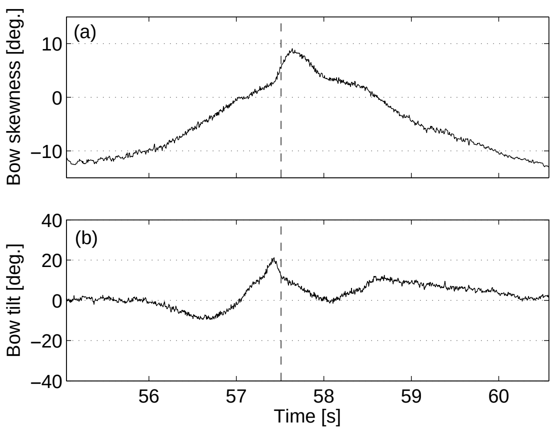


Figure 5: Bow angles during crescendo-decrescendo (up-down): (a) skewness, and (b) bow tilt

pressure is exerted by a torque in the pitch direction, mediated by the index and pinky fingers. This torque is associated with the compliance of the bow and the string experienced by the player when pressing the bow into the string. This lever mechanism means that the player can exert less force toward the tip of the bow. Consequently, considering the examples in section 2.2, a crescendo will come more naturally for the player during an up bow. Additionally, as described and implemented by Nichols [5], the pitch axis in combination with the vertical direction also allows the player to feel the location of the strings (detents).

We saw that the presence of skewness in the bow's yaw orientation creates pull on the bow along the string, felt as a torque by the player, causing the hair-string interaction point to drift. After change of bowing direction, this drift will also occur in the opposite direction. However, it is often the case that after some amount of horizontal movement, the player will preemptively counter this drift by applying an opposing torque. Furthermore, when bowing straight the contact with the string keeps the bow from moving in lateral direction.

In cases where the player has employed bow tilt, downward force is accompanied by a slight torque in the roll direction, created due to the stick-hair orientation no longer being in parallel with gravity. Changes in tilt also affect the number of hairs in contact with the string, which in turn changes the vertical spring constant as well as the frictional resistance to orthogonal movement across the strings.

4 Conclusions

We have reviewed some aspects of control gestures in violin bowing and described observations of their use in co-incidence with changes to sound-related bowing parameters. We have also described some aspects of force feedback experienced by the player in relation to these control gestures.

Previously, apparatus supporting force in two [6], three [3], and four [5] degrees of freedom have been used for implementing virtual bowing interactions. We suggest that this is not sufficient for a complete simulation of forces experienced by the player; the above analysis shows that torque is felt in all three rotational directions: pitch, to feel string detents and to apply vertical pressure; yaw, in relation to changes of bow-bridge distance; and roll, when tilt is employed. Combined with two of the three linear directions (frictional force and compensation for gravity), it becomes clear that it is necessary to employ force in at least five axes when synthesizing a virtual haptic representation of the bowing gesture, with freedom of movement in the lateral direction. Future experiments with 6-DOF force feedback devices are planned to assess the relative importance of the forces and torques experienced by players.

Acknowledgments

Special thanks go to Matthias Demoucron for developing and sharing the bow force sensor. This work was sponsored in part by the Natural Sciences and Engineering Research Council of Canada and the Enactive European Network.

References

[1] A. Askenfelt. Measurement of the bowing parameters in violin playing. II: Bow-bridge distance, dynamic range, and limits of bow force. *J. Acoust. Soc. Am.*, 86(2):503–516, August 1989.

[2] C. Cadoz. Instrumental gesture and musical composition. In *Proceedings of the 14th International Computer Music Conference*, pages 1–12, Köln, 1988. ICMA.

[3] J.-L. Florens. Expressive bowing on a virtual string instrument. In A. Camurri and G. Volpe, editors, *Lecture Notes in Computer Science: Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop*, volume 2915, pages 487–496. Springer Berlin / Heidelberg, Genova, Italy, April 2003.

[4] K. Guettler, E. Schoonderwaldt, and A. Askenfelt. Bow speed or position – which one influences spectrum the most? In *Proceedings of Stockholm Music Acoustics Conference*, pages 67–70, Stockholm, 2003.

[5] C. Nichols. The vBow: development of a virtual violin bow haptic human-computer interface. In *Proceedings of the Conference on New Interfaces for Musical Expression*, pages 29–32, Dublin, Ireland, 2002.

[6] S. M. O'Modhrain. *Playing by Feel: Incorporating Haptic Feedback into Computer-Based musical Instruments*. PhD thesis, Stanford University, November 2000.

[7] J. C. Schelleng. The bowed string and the player. *J. Acoust. Soc. Am.*, 53(1):26–41, 1973.

[8] E. Schoonderwaldt, K. Guettler, and A. Askenfelt. Effect of the width of the bow hair on the violin string spectrum. In *Proceedings of the Stockholm Music Acoustics Conference*, pages 91–94, Stockholm, 2003.

Engagement and Immersion Research in Interactive Art Environment: exploring subjective and physiological data based on different visual cues

Jinsil Seo
Simon Fraser University, Canada
E-mail: jinsils@sfu.ca

Abstract

This research aims to understand what influences experience in a projected immersive environment. In this paper, I am specifically going to examine if one of strong visual attentions, motion, affects to create engagement and how engagement correlates to immersion and vice versa. To investigate this, physiological data was collected from a small sample of participants as well as questionnaires and interview data, and analyzed. The experiment revealed that participants became more engaged by moving star images and their engagement level correlated directly with the speed of stars. Through comparative analysis of the physiological data, it was found that the static image sessions caused faster heart rates; and, as long as the images moved, the average heart rates got slower. In addition, the biofeedback data showed an inverse correlation between animation speed and GSR level. In terms of immersive experience, people tended to be more immersed with faster animation.

1. Introduction

Immersion comes etymologically from the Latin word, in- "into" + mergere "plunge, dip." It has evolved to mean "absorption in some interest or situation." As a method of teaching a foreign language, it is from 1965, trademarked by the Berlitz Company. [1]

The technology to immerse people in computer generated worlds was also proposed by Sutherland in 1965, and realized in 1968 with a head-mounted display that could present a user with a stereoscopic 3D view slaved to a sensing device tracking the user's head movements. [2]

Immersion in VR has been described as a technology, which can be achieved to varying degrees. The degree of immersion is increased by adding "consistent modalities, greater degree of body tracking, richer body representations, decreased lag between body movements and resulting changes in sensory data, and so on…" [3]

Since 2002, I have been working on immersive responsive installations focusing on physical elements mainly touching. To create interaction, I have tried projected images, fiber optics, and LEDs for my installation. While working with them, I began wondering how visual cues (sometimes having physical interfaces as in touch screens) affect people's experience in immersive space. In addition, what is that defines the immersive experience? I focused on creating immersive experience in different ways (a more artistic way) from other scholars who used VR, narratives, etc.

As a pilot study, this research aims to understand what influences experience in a projected immersive environment. I am focusing on engagement and immersion despite knowing rationally that people will see and interact only with projected images. In this paper, I am going to examine how visual attentions (cues) affect engagement and how engagement correlates to immersion and vice versa. To investigate this, physiological data was collected from a small sample of participants, and analyzed with other method data.

Immersive and engaged experience is the ideal of affective experience. [4] Through these experiences, people can find pleasure and meaning for themselves. It will provoke to create more internal and touched experience.

2. Background Research

2.1. Motion as a visual cue in visualization

According to *Lyn Bartram*, motion has compelling properties which make it potentially very useful in designing user interfaces. Motion holds promise as a perceptually rich and efficient display dimension. She examined how motion enhances visualizations and tried empirical studies. Among many visual cues for visualization: color, contrast, etc., motion becomes increasingly important with large displays and complex information areas.

2.2. Research on immersion

Mel Slater is a VR researcher focusing on immersion and presence in VR immersive environments. He distinguishes between immersion and presence. Immersion includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching. He stated presence is a state of consciousness, the psychological sense of being in the virtual environment. Participants who are highly present should experience the virtual environment as a more engaging reality than the surrounding world, and consider the environment specified by the displays as places visited rather than as images seen. Even though he constraint the concept of immersion and presence to only consider Virtual Reality, Slater's research offers a background knowledge of immersion and could be extended and implied in artistic ways [5]

Janet Murray talked about immersion in new media, especially interactive narrative media. In her book Hamlet on the Holodeck, she stated that immersion arises from our pleasure in complex, extensive, consistent alternate worlds that enfold us while still allowing us to explore. [6]

2.3. Measuring subjectivity with biofeedback

Rosalind Picard is a pioneer researcher into emotion and affective computation. Her main research is creating new techniques to assess frustration, stress, and mood indirectly, through natural interaction and conversation. For her research, Picard often uses physiological measurement to recognize what people's emotion states are. [7]

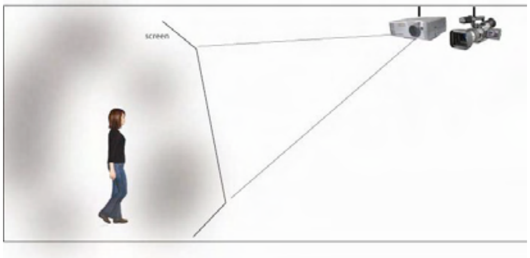
Meehan, et al found that experience in stressful virtual environment would evoke physiological responses similar to those evoked by the corresponding real environment, and that greater presence would evoke a greater response. To examine this, they conducted three experiments, the results of which support the use of physiological reaction as a reliable, valid, sensitive, and objective presence measure. The experiments compared participants' physiological reactions to a non-threatening virtual room and their reactions to a stressful virtual height situation. [8]

Regan, et al (2005) presented that a method of modeling user emotional state, based on a user's physiology, for users interacting with play technologies. They measured GSR, EKG, EMG, respiration and showed a great potential for using physiological metrics to model emotional experience with interactive play techniques. [9]

3. Experiment

For this experiment, participants were not given definitions of engagement and immersion so that they judged these qualities based on their own criteria and understanding. There were 12 participants: 10 males and 2 females. The immersive environment was set up like figure 1 using two black fabric walls (2.5m*3.5m) and one curved white fabric wall, a LCD projector, and a video camera.

Figure 1. Illustration of the Setting



Participants were fitted with Thought Technology Ltd. FlexComp+ EKG sensors on their chest and GSR sensors on their non-dominant hand. The software for visualizing and recording these measures was run on the same PC with the visualizing system. In the environment, star images were employed with three-levels of different states: static, slow animation, and fast animation. I used star images because in the hope that it would offer people a more immersive sense by reminding them of watching the sky at night surrounded by stars. Stars also are a feature in my earlier artistic experiments. During the experiment, participants were shown three different images for two minutes each session. *Session 1 (Static)* was a realistic star photograph. People were free to move and walk around the space during the session. In *Session 2 (Slow),* Stars were animated (Figure2). Stars moved very slowly following people's shadow on the screen. Depending on the darkness of people's shadows, the stars' sizes and brightness were being changed. Dark shadows created bigger and brighter star movements. *Session 3 (Fast)* was same interaction as session 2 except faster speed. Data were collected based on questionnaires, interview, and physiological data.

Figure 2. Screenshot of fast animation

3. Data Collection

Data was collected from four different sources: Questionnaire, Interview, Video recording, Biofeedback data (GSR and EKG). People were asked if they did get engaged after each session and they marked in the questionnaire (How much did you get engaged in this session?). This was measured on a 7 point Likert Scale, from 1 = “Not at all” to 7= “very much”.

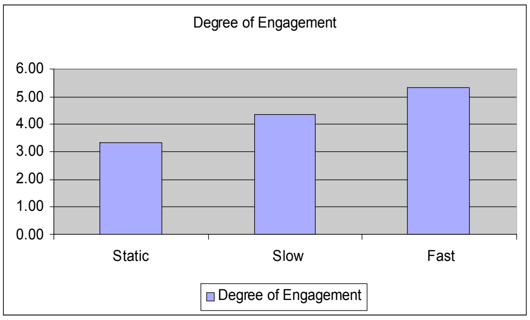
Physiological data is not available for all participants because the equipment did not always function correctly for the duration of the session due to body motion or shifting in the sensor positions. Data was collected while each participant was in the experiment setting.

4. Result

4.1. Engagement

Participants rated the engagement on 7 point scale after seeing each animation: static, slow and fast. The mean results are shown in table 1(static mean = 3.3, slow mean = 4.3, fast mean = 5.3). According to the table 1, people became more engaged by moving star images and became more engaged by speed of stars.

Table 1. Degree of Engagement



In terms of physiological data analysis, in the static images sessions, people’s heart rates were faster. As long as the images moved, the average heart rates got slower. This pattern is distributed across all participants in Table 2; however, each person didn’t have same experience as the average value (meanHR_static = 98.62, slow = 79.58, fast = 78.55). To correlate the questionnaire subject ratings to the physiological data, I would say that of the visual cues, increased animation correlated to increased engagement and speed also correlated to increased engagement in the immersive environment.

Table 2. Average of HeartRate

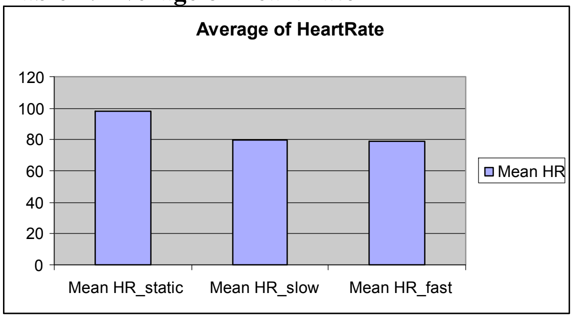
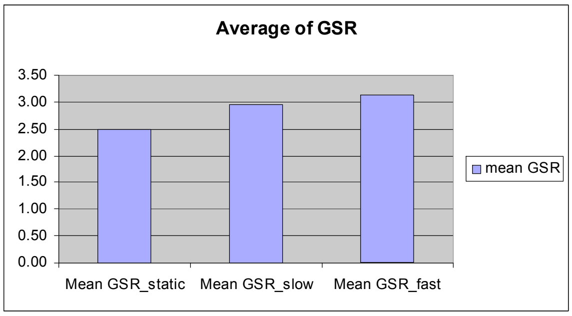


Table 3. Average of GSR



The participants’ Galvanic Skin Response data also showed an opposite pattern in Table 3. Their GSRs increased based on animation and the animation’s speed (mean GSR_static = 2.50, slow = 2.96, fast = 3.12). Comparing questionnaires data and physiological data, It could be concluded that when a person becomes engaged, heart rate becomes slower, but GSR becomes faster.

4.2. Immersion

After the session, each participant answered a question: Did you feel immersion during the experiment? And after filling out the questionnaire, they had an interview and explained during what session they felt immersion, and how the experience was. 11 participants out of 12 answered that they felt immersion during at least one of the sessions. But sessions they felt immersion in, were quite different as table 4 below.

Table 4. Number of people who got immersed in each stage.

Static	Slow	Fast
4(36%)	2(18%)	5(45%)

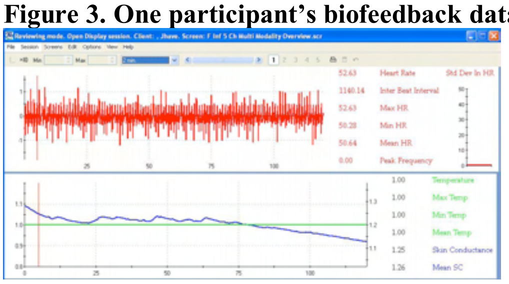
Though 45 percent of participants felt immersion in the fast animation session, since the sample size was relatively small, the real numerical difference between participants immersed in static versus fast image is only one. So it is difficult to conclude that people got easier immersed within the fast responsive animation. Based on each participant's subjective understanding of the concept of immersion, the point at which they felt immersion was different.

During the interviews, some of participants who have a lot of experience of VR and immersive environment said. "I felt immersion in the static session, because that static image looked real like a photographic image. I believe that to get immersion, you better try with realistic images, maybe even better with 3D." Some other people who answered "static" for the session they felt immersion articulated that they saw more depth in the photo. That's why they felt immersion with the static image. The animated images didn't have depth; they looked very flat and lacked depth. In understanding immersion, depth perception is an important visual cue.

From the video analysis, some people tended to move actively and some other people tended not to move a lot when they got immersed.

One of the subjects said the slow animation was the most immersive because that was new and slow animation creates thorough gesture movement. His GSR showed in Figure 5. In the graph, a blue line presents that GSR value gets decreased as long as the experiment time goes. Correlating to his physiological data, immersion made GSR slow down.

Figure 3. One participant's biofeedback data



5. Conclusion and Future Work

In this paper, I aimed to focus on if motion influences people's immersive and engaged experience. The experiment revealed that participants became more engaged by moving star images and their engagement level correlated directly with the speed of stars. Through comparative analysis of the physiological data, it was found that the static image sessions caused faster heart rates; and, as long as the images moved, the average heart rates got slower. In addition, the biofeedback data showed an inverse correlation between animation speed and GSR level. In terms of immersive experience, people tended to be more immersed with faster animation. However, it is difficult to conclude that people got more easily immersed within the fast responsive animation. Based on each participant's subjective understanding of the concept of immersion, it is clear that the point at which they felt immersion was different. For future work, I would like to have more subjects to test the validity of the experiment. Finally, for the sense of immersion, sound is also one very important aspect. So I believe that surround sound will create a richer immersive experience.

5. Acknowledgement

Thanks to Professor Diane Gromala for her strong academic passion and encouragement, and obtaining materials and a space for this research.

References

[1] Online etymology dictionary: http://www.etymonline.com/
[2] Slater, M., Steed, A., McCarthy, J. and Maringelli, F. (1998) The influence of body movement on subjective presence in virtual environments, Human Factors 40(3), 469-477
[3] Virtual Reality Introduction & Definition : http://vrlab.epfl.ch/~fvexo/VR_2007/slides/EPFL-Virtual%20Reality-FV-2007-Introduction_Definition.pdf
[4] Douglas Y., Hargadon A. (2000) The pleasure principle: immersion, engagement, flow, Conference on Hypertext and Hypermedia, Proceedings of the eleventh ACM on hypertext and hypermedia, 153-160.
[5] Slater, M., Steed, A., McCarthy, J. and Maringelli, F. (1998) The influence of body movement on subjective presence in virtual environments, Human Factors 40(3), 469-477
[6] Murray, J. H. (1997) Hamlet on the Holodeck: The Future of Narrative in Cyberspace, The Free Press, New York, NY.
[7] Picard, R. W. (1997) Affective computing, MIT Press, Cambridge, MA.
[8] Meehan, M. J. (2001) Physiological reaction as an objective measure of presence in virtual environment, Unpublished doctoral thesis, University of North Carolina, Chapel Hill.
[9] Mandryk, R. L., Atkins, M. S. and Inkpen, K. M. (2006) A Continuous and Objective Evaluation of Emotional Experience with Interactive Play Environments. Proceedings of the Conference on Human Factors in Computing Systems (CHI 2006). Montreal, Canada, April 2006, pg. 1027-1036.

The Meditation Chamber: Enacting Autonomic Senses

Chris D. Shaw, Diane Gromala,
School of Interactive Arts and Technology,
Simon Fraser University, Canada

A. Fleming Seay
School of Information,
University of Texas, USA

E-mail: shaw at sfu.ca, gromala at sfu.ca, fleming_seay at dell.com

Abstract

The Meditation Chamber is a novel combination of immersive virtual reality (VR) and biofeedback technologies. Its primary goal is to help users lower their stress levels through meditation and muscle relaxation. Its secondary goal is to enable users who have never meditated to gain real-time feedback that allows them to sense when they are indeed lowering the physiological states that are indicative of meditating. During its first use – an exhibition at SIGGRAPH 2001, between Emerging Technologies and the Art Gallery — the relaxation levels of 411 users were studied. It is now in use at Virtually Better and its clinical partners. This paper explores its first iteration, and outlines the plan for its reconfigured, second use for the treatment of chronic pain.

1. Introduction

This paper describes the design and analysis of the *Meditation Chamber*, a bio-interactive, therapeutic, virtual environment. It was initially exhibited at SIGGRAPH 2001 and after minor refinements, is now in use at Virtually Better,[1] along with its partners.

The goal was to design and build an immersive virtual environment that used visual, audio, and tactile cues to create, guide, and maintain a user's guided relaxation and meditation experience. Real-time biometric data was used to partially control most of the environment, thus shaping the user's experience. The methodology for this project combined objective and subjective data collected from the participation of over 400 attendees at SIGGRAPH 2001. The subjective methodology included a phenomenological analysis of users and the VR artists.

2. The Effectiveness of Meditation

It is estimated that meditation has been systematically practiced at least for a millennium [3]. In the last decade, meditation has assumed legitimacy in Complementary and Alternative Medicine and Integrated Medicine in North America and some European countries. Scientific studies have largely affirmed its benefits, including its psychological and physical efficacy and claims [1,6,8,9,10]. Meditation practices, such as those by Dr. Kabat-Zinn, have made significant inroads to the acceptance of meditation in Western normative medical practices. It can be argued that bio-feedback achieves similar abilities. Miller points out that biofeedback ". . . can be used to help people to learn to improve the perception of certain visceral events" [4]. Both meditation and biofeedback enable users to bring into conscious awareness those physiological states that usually remain in the background of awareness.

We chose the Thought Technologies ProComp+ for biofeedback because of its ease of use and ergonomic efficiencies. We chose to integrate this with an immersive virtual environment, in order to enable users to get real-time feedback and to gain a sense of control over their autonomic senses. More accurate measures of a user's subjective state can be had with EEG, but these devices take a long time to put on, and require careful calibration. An adequate indication of changes in real-time states was indicated with our use of respiration, heart rate, and galvanic skin response. The authors follow Rosalind Picard's sense that biofeedback devices are only suggestive of changes in states [5], and are not able to tell us when a user is in a meditative state with any absolute certainty.

The purpose of Buddhist practices of mindful/awareness "is to become mindful, to experience what one's mind is doing as it does it, to be present with one's mind" [7]. Since this takes a great deal of regular practice, the most we could hope for in the short time of testing the *Meditation Chamber* was to introduce users – particularly novices – to their abilities to perceive and control some of their autonomic senses within a visual and sonic environment.

Awareness of interoceptive sensation usually makes itself known when serious phenomena, like food poisoning or cardiac arrest are imminent. Our initial concern was with stress, which affects almost

[1] Virtually Better is a treatment facility in Atlanta, Georgia that uses and develops VR technology for psychotherapeutic purposes.

every part of the visceral systems. The seriousness of stress is often dismissed, underrated or misunderstood. Stress has extremely negative consequences: it negatively affects immunity and endocrine production and is "linked to the six leading causes of death: heart disease, cancer, lung ailments, accidents, cirrhosis of the liver, and suicide" (Canadian Mental Health Association 2007). Thus, alternative therapeutic techniques related to relaxation and the management of stress are increasingly employed to augment traditional treatment by drug-based, medical therapies.

3. The Mediation Chamber

For people who are new to the practice of any form of meditation, one of the difficulties is in knowing if or when their efforts are indeed achieving anything. The resultant *Meditation Chamber* is comprised of an immersive virtual environment, a biofeedback device, and customized programming using the SVE Toolkit.

Figure 1: The Meditation Chamber

One roadblock discovered to impede the effectiveness of relaxation therapies is the consistency and quality of the user's experience. Most formally trained doctors or assistants are not knowledgeable enough to administer alternative treatments. Also, some people have difficulty with visual imagery and are not good candidates for meditation exercises. The goal of the research project presented here was to design and build an immersive virtual environment that used visual and audio cues to create, guide, and maintain a user's guided relaxation and meditation experience.

There are several possible advantages to using a virtual environment to support meditation and guided relaxation beyond providing new meditators with real-time feedback. Patients without good mental imaging skills would still be able to benefit from the use of meditation. Clinicians with minimal training in meditation and guided imagery would be able to provide a consistent, high quality relaxation/meditation experience to their patients. Also, by providing specific meditation environments, we can guarantee that participants in future studies all receive identical training and treatment. The use of meditation and guided imagery is well established for its utility in the treatment and prevention of a number of several diseases with high cost in terms of both human suffering and financial cost. The possibility of increasing the effectiveness and repeatability of this type of therapy may receive a great deal of interest from the medical community. To date, we have received numerous inquiries.

4. System Design

Galvanic skin response, respiration, and blood volume pulse were chosen as the biometrics of interest for this project. These biometrics were collected using the ProComp+, a commercially available device produced by Thought Technologies. The sensors described below are all standard biometric sensors produced by Thought Technologies for use with the device. Galvanic skin response (GSR), commonly used in lie-detector tests, is a measure of the change in the electrical conductivity of the skin that results from the body's reaction to emotional stimuli. It is fairly useful in measuring an individual's general level of arousal as well as tracking changes in arousal as they relate to events in the individual's environment.

GSR is measured by attaching two electrodes to the user's fingertips and measuring conductivity changes in a reference charge passed through the user's skin. The reference charge is weak enough so that no sensation is created.

Respiration rate was measured using a flexible chest strap that was stretched around the user's upper chest and fitted just below the armpits. The strap was equipped with a length of rubber tubing that flexed and relaxed as the user's chest expanded and contracted during respiration. Heart Rate was measured using a blood volume pulse sensor that monitored cardiac pulse at the tip of the index finger. The audio-visual content of the environment was delivered to the user via head mounted display.

The head mounted display used for this installation was the VFX-3D, produced by Interactive Imaging Systems. This bi-ocular HMD does not have a stereo-visual display, but gives the user an approximately 60-degree field of view, which is larger than most. The unit's large, high quality headphones also figured in to its selection for their ability to deliver robust sound and cut down on the intrusion of external noise. A library written by Thought Technologies allowed us real-time access to the stream of data produced by the sensors. We were then able to write code that allowed us to use this data stream to manipulate aspects of the environment. This feedback loop let the user's bio-rhythmic state alter the environment in subtle ways just as the environment worked to relax the user.

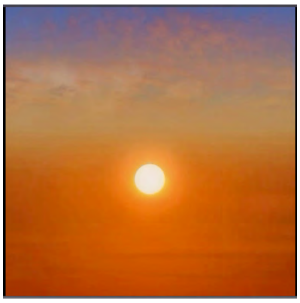


Figure 2: Image of the sunset taken from initial phase.

5. Content

This initial phase served to relax users and introduce them to the experience. After being asked to breathe deeply and relax, the user was presented with a visual depiction of the sun just before sunset. A narrator's voice told users that the sun would descend in the sky as they relaxed, breathed deeply, and flushed their mind of worldly concerns (see Figure 2). The speed of the sun's descent depended on the rate of GSR decrease. The second part of this relaxation phase operated in the same way as the first, but depicted a moonrise instead of a sunset.

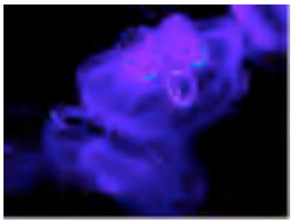
At each one second measurement interval, the average current GSR reading was compared to the previous second's GSR reading, and if the GSR value had increased over the time period, the current frame of the sunset/moonrise animation would be paused. If there was a decrease in average GSR, the sunset/moonrise animation would step forward one frame. Depending on the frame-rate the user achieved, this combined sequence took 2-4 minutes to complete. Users were not told explicitly of the relationship between their GSR and the frame rate of the animation so that they would make no effort to "play" the environment and would instead concentrate on the experience of relaxation.

The second phase of the experience was a guided, progressive muscle relaxation exercise. The user was coached to flex, hold, and release a set of eight different muscle groups including the legs, arms, abdominals, and shoulders (see Figure 3). Each muscle group sequence was accompanied by gender appropriate visuals depicting the described motion, usually from a first person perspective. Based on user preference, male users viewed a male body performing the exercises while female users viewed a female body. This phase was not initially interactive, but instead users listened to the narrator's instructions while mimicking the movement examples visually presented to them on the screen. The progressive muscle relaxation phase lasted roughly 6-7 minutes. The nature of flexing and relaxing the major muscle groups provided a tangible experience of the creation and, more importantly, the release of tension.

The third phase was designed to teach users a basic meditation called "following your breath," in which users are asked to focus all their awareness on the sensation of their breath coming and going from their nostrils. If other thoughts entered their awareness during this time, they were told to push them aside calmly and firmly and to remain focused only on their breath. This phase lasted approximately seven minutes and was accompanied by an abstract visual display created by putting several image filters on top of video of a swimming jellyfish.

Figure 3: Muscle relaxation. Phase two, the muscle relaxation phase, was guided by visuals and audio narration.

The image seems to pulse in time with users' respiration. The audio during this segment sounded like abstracted, calmly moving water and was sampled from sounds taken from a waterfall. The jellyfish images faded and disintegrated as users' biofeedback measures were reduced, and eventually faded to black.



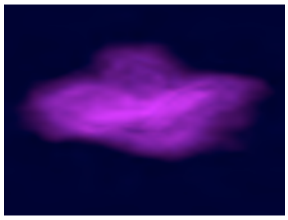


Figure 4: In Phase 3, two abstract jellyfish images.

6. The Installation

The system was installed at the Emerging Technologies Exhibition at SIGGRAPH 2001, with four stations as described above. When users arrived at their reservation time, they completed a brief pre-experience inventory, and were briefed on the experience. Users were told that their biorhythms would control the environment in subtle ways, but they were not told specifically how this would occur. Users were then seated in one of the private booths and fitted with the sensors and the HMD. When ready, the application was started and the user was left alone in the booth. At the end of the experience, each user completed a post-experience inventory asking how relaxed s/he felt following the experience, and soliciting any comments.

Each user's biometrics (GSR and respiration) were printed out as a line graph on the feedback sheet, which contained an explanation of how to analyze the graphical representations of the experience. Almost every user took their print-out, and approximately 85 percent of users remained to get more information.

7. Results

411 SIGGRAPH 2001 attendees experienced the installation during the five-day exhibition. We can report on the completed analysis of the subjective measures of relaxation collected in order to assess how effective the installation was in providing each individual with a relaxing experience.

We have analyzed the extensive amount of biometric data collected from the SIGGRAPH attendees, concentrating mostly on the GSR data. Analyses show that two general patterns of GSR profile can account for nearly 75% of the generated data and are each generally accompanied by two distinct respiration profiles. Just over half of the participants exhibit what can be called a “novice” GSR profile. This means that their GSR level starts relatively high, descends through the first phase of the experience, increases and shows peaks in the muscle relaxation phase, and then begins to decline again in the final phase, ending up at or usually beneath the low established in the first phase.

Breathing patterns in individuals exhibiting the novice profile tend to be steadier and deeper in the final phase than in the first phase. The second profile, which accounts for nearly a quarter of the GSR data, is termed the “expert profile.” Individuals exhibiting this profile show precipitous drops in GSR during the first phase, entering a very low and often flat GSR state before the muscle relaxation phase begins. This flat-line state is typically maintained throughout the remaining two phases, and is accompanied by a very steady but not necessarily deep breathing pattern. Individuals exhibiting the expert GSR profile also show very consistent respiration rate and amplitude through-out the experience.

We later tested a baseline condition with a user sitting quietly in a room wearing the biofeedback hardware for 20 minutes. Of our 16 baseline subjects, 8 experienced increases in GSR over the 20 minute time period. This indicates that the Meditation Chamber is effective at promoting relaxation.

Users self-rate rated their level of relaxation from 1 (very anxious) to 10 (very relaxed). The average pre-session relaxation rating was 5.63, with a Standard Deviation (SD) of 1.75. The average post-session relaxation rating was 8.00, with a Standard Deviation of 1.69. A t-test showed that post-session relaxation ratings ($M=8.00$, $SD=1.69$) were significantly higher than pre-session ratings ($M=5.63$, $SD=1.75$), $t(410) = -24.45$, $p=.0001$.

Most of the subjective comments were positive. Negative comments had to do with the heaviness of the HMD, and the noise from the other exhibits. The positive comments were that users felt relaxed, though many initially wrote that they did not expect to. At exit, when we gave users a copy of their changes in states, they became even more positive and expressed surprise that they could “learn” how to lower their measures and “feel” them.

Subjective methods also included first-person, phenomenological studies, which resulted in a dissertation and several articles. These studies were integrated with objective measures have significantly impacted the reconfiguration that is underway for use in chronic pain patients.

Individuals who suffer from chronic pain face many more stressors than those who suffer only more acute pain episodes. Physicians in North America understand chronic pain in a biopsychosocial model [2]. While the biological processes of chronic pain itself cause the body profound stress, psychological and social factors often magnify this stress. Meditation, though not a cure, has been found to lower stress levels and thus put a potential brake on the spiraling of biopsychosocial stressors. The Meditation Chamber, in helping new users meditate by providing feedback, will be tested in this realm for its potential benefits.

9. Future Work

Essentially, the artistic or representational aspects are being redesigned to better accommodate cultural factors, and to integrate the subjective aspects of Buddhist meditation. Thus, users will be able to chose from “medical visualizations,” “Traditional Buddhist” imagery and sound, “Comparative Medicine Histories,” and “Abstract” representations. In addition, the sequence of meditation and guided muscle relaxation are reordered, and sensors attached to user’s knees and wrists will sync up with the muscle relaxation graphics.

These extensions to the Meditation Chamber have the proven potential to enhance a user’s ability to learn how to meditate, and in doing so, reduce the damaging effects of stress.

References

[1] Barnes,V., Schneider, R., et al. 1997. “Stress, Stress Reduction, and Hypertension in African Americans: An Updated Review,” *J. Natl Med. Assn,* 89(7):464-476.

[2] Gatchel, R.J., Peng, Y.B., et al., 2007. “The Biopsycho-social Approach to Chronic Pain: Scientific Advances and Future Directions”, *Psychological Bulletin* 133(4):581-634.

[3] Gethin, R., 1998. *The Foundations of Buddhism* (OPUS), Oxford: Oxford University Press, New York, pg 1.

[4] Miller, N.E., 1977. “Biofeedback and Visceral Learning.” In *Biofeedback and Self-Control*. Stoyva, J, et al., eds. New York: Aldine.

[5] Picard, R., 1997. *Affective Computing,* Cambridge, MA: MIT Press. 122

[6] Segal Z., Williams, et al. 2001. *Mindfulness-Based Cognitive Therapy for Depression: A New Approach to Preventing Relapse,* Guilford Press, New York.

[7] Varela, F.J., Thompson, E., et al. 1993. *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge MA: MITPress.

[8] Walton K.G., Pugh, N.D., et al. 1995. “Stress Reduction and preventing hypertension: preliminary support for a psychoneuroendocrine mechanism.” In *J. Alternative Complementary Medicine,* 1(3), 263-283.

[9] Wenneberg, S.R., Schneider, R.H., et al. 1997. ”Anger expression correlates with platelet aggregation,” *Behavioral Medicine*, vol. 22, no. 4, pp.174-177.

[10] Zamarra, J.W., Schneider, R.H., et al. 1996. “Usefulness of the transcendental mediation program in the treatment of patients with coronary artery disease.” *Am.J. Cardiology,* 77(10), 867-870.

An Attempt of Dance Sequence Composition
- Blending Ballet and Contemporary Dance -

Asako Soga* Bin Umino†
(*)Ryukoku University, Japan
(†)Toyo University, Japan
E-mail: asako@rins.ryukoku.ac.jp, umino@toyonet.toyo.ac.jp

Abstract

The purpose of this study is to develop a program which allows users to create and simulate blended choreographies of classical ballet and contemporary dance. We try an interactive composition and an automatic composition of dance sequences. The blended choreographies are derived from our basic concept that is to separate a dance movement into elemental motions and use them like building blocks. The experiment with a professional choreographer and dancers shows us that the system is useful for both choreographers and dancers by inspiring and suggesting them novel choreographic idea.

1. Introduction

There have been many studies using human motion data such as controlling parameters of motions [1], blending some resources [2], and synchronizing to music [3]. In dance field, there have been many interactive performances using multimedia data [4]. In recent years, 3-dimensional motion data is often used in performances because it is getting easy to record human motion by using a motion capture system.

Our goal is to develop useful tools in dance education and creation such as a self-study system for students and a creation-support system for artists. Our research approach focuses on creating and composing new choreographies of dance.

In 1980's, Merce Cunningham used the computer software called LifeForms [5] to discover his original movement. This software is the most famous application that can compose and edit dance scores. In recent works [6], it has been extended to computer notation systems by using Labanotation [7], which is one of the most famous dance notations. By using this system, the user can simulate strictly scored dance animation. However, it is difficult to compose creative and effective choreographies.

Our basic concept is "to separate dance movement into elemental motions and use them like building blocks." We call it "analytic-synthetic composition system." Our strategy is completely different from Cunningham's method, but it generally takes over the two ideas from Cunningham. Constraint and contingency of choreography are another strategy of finding novel movement.

We have already developed an interactive composition system for ballet called "Web3D Dance Composer (WDC)" [8]. The system allows for users to simulate dance steps and step sequences in 3D animation on the Web. The system has possibility of "e-learning" for ballet, and it may be helpful for computational dance research. We shift our target from educational purpose to creative phase.

One of the contributions of this work is to develop a composing system that can be used for all kinds of dance. In this research, we try to create sequences by blending ballet steps and contemporary movements. Ballet encompasses the most basic movements and its method is strictly defined. The method complies with esthetics of formal beauty of ballet which has been brought up for 400 years. On the other side, contemporary dance has no basic steps and no limit of motion. Therefore, the proposed system can be adapted for any other dance genres.

This research is a feasibility study of blending classical ballet steps and contemporary dance movements. Our proposed systems have adapted for ballet and contemporary dance. We intend to evaluate how useful these systems are for creation of a new dance work.

2. Interactive composition of blending ballet and contemporary dance

First we have developed an interactive composition system for ballet and contemporary dance. Then we conducted an evaluation test by a professional choreographer who is also a dancer.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

2.1. Archiving dance motions

We constructed a “Dance motion archive” of ballet and contemporary dance. To make the dance motion archive, we divided dance movement into ‘manipulated units,' that is the short pieces of motion, for the benefit of information storage and retrieval. The basic steps and movements were performed by professional dancers using motion capture systems to gather 3D motion data.

For archiving ballet steps, we recorded most natural arm movements that correspond to each step at the same time because most steps represented the lower body movements.

In the case of contemporary dance, there are no standard steps and no typical methods that have to be considered. Thus, a professional choreographer created 29 movements which were characteristic for contemporary dance. Table 1 shows the examples of contemporary dance movements. We recorded these as the basic movements for contemporary dance.

In this prototype system, it includes 305 movements. It consists of 247 basic ballet steps and 58 typical contemporary movements.

2.2. Constraints for interactive compositions

We have expanded an interactive composition system for ballet into the system for contemporary dance. There are many strong constraints in our proposed system for ballet. Each step should be selected so that the ending pose of the previous step becomes the beginning pose of the following step. In order to adapt the contemporary dance, the starting pose and ending pose of every contemporary dance movement is with the classical ballet foot position, such as 1st, 4th, and 5th position. Then the motion catalog of all steps/movements is classified by the starting pose. Figure 1 is the user interface of the WDC for ballet and contemporary dance.

2.3. Experiment

We conducted an experiment on the interactive composition system to evaluate its practicality. A professional contemporary choreographer who knows ballet well and also a dancer evaluated the system.

First she created 7 sequences of contemporary dance by using interactive composition system. Our request was to make somewhat novel sequences which last about 20 seconds. After previewing each sequence by 3D animation a few times, she performed two types of dance: “Representational performance of 3D simulation” and “Arranged performance of the 3D simulation.” The arranged performance is also the representational performance, but added some slight arrangement.

Table 1. Examples of contemporary movements

Name	Explanation of the movement
neck-roll	Rolling the neck
neck-side	Shaking the neck to side by side
hip-roll	Rolling the chest
arms-swing	Swinging the arms
leg-shake	Shaking one leg
balance-off	Break off the body balance
floor-acro	Acrobatic movement on the floor

Figure 1. User interface of WDC

Figure 2. Performance scenes of the creative sequence

Then we asked her to create more creative sequence which lasts about 1 minute after selecting one of 7 sequences. This is “Experimental performance inspired by the 3D simulation.” Figure 2 is the performance scenes of the creative sequence.

After the experiment, she gave us several important comments from the point of a professional choreographer.

2.4. Results and discussion

She pointed that there is very long distance between the seeds the system affords and the real creative performance on the stage. What fills in the gap is intelligence of the choreographer and body sensation of the dancers. At the experiment, she said that the dancer needs to let their body sing or let the body take a trip, if you wanted to approach the creative work from the 3D animation. If the system succeeds in planting seeds, an experienced dancer can let the seeds grow.

She also pointed that from a contemporary standpoint, the classical ballet foot position, such as 4th, and 5th position, is too strong constraint for choreography. At the moment, every movement starts with the position and end with the position. This constraint make the system easily connect movements in 3D environment. However, as the body is forced to return to the position each time, the sequence is always prevented to flow smoothly and fluently.

The system allows a user easily create un-natural series of movement of human body. In many cases, they can be undesirable noise for choreography. But sometimes they can be unexpected clues.

3. Automatic composition for ballet and contemporary dance

We have developed an automatic composition system for ballet and contemporary dance. Then we conducted an experiment of automatic composition.

3.1. Concept and basic algorithm

The algorithm of composition built in the program can automatically compose the ballet step sequence using the step archive. The basic approach of the composition method is selecting each step randomly from the beginning of a sequence to the end.

The automatic composition algorithm for ballet has many strong constraints. To adapt the contemporary dance, we aim to make a rough list of basic movements and not to make complete animations because a dancer who knows about contemporary dance can recognize the sequence and arrange it freely. Therefore, we do not need to consider the strict poses when connecting steps or movements. We skipped and reduced the step arrangement of ballet because dancers could arrange basic steps/movements at the beginning or ending time. In this prototype system, it includes 118 movements. It consists of 60 basic ballet steps and 58 typical contemporary movements.

3.2. Arrangements of choreographies

Contingency of choreography is another strategy of finding novel movement. As the way to arrange choreographies, we suggest changing the timing of some movements. Therefore, we have implemented the beat change of each step/movement. The prepared beat speed are "fast", "normal", and "slow." The corresponding beat are "-1 beat", "0 beat", and "+1 beat" to the original beat. In a case when the original beat is under 1 beat, the "fast" speed should be a half beat.

Figure 3. A scene of the experiment

3.3. Experiment

We conducted an experiment on the automatic composition system to evaluate its practicality. Three professional contemporary dancers who well know ballet evaluated the ten sequences. Figure 3 is one of the scenes of the experiment.

After previewing each sequence by 3D animation a few times, they performed the representational choreographies. For performing the choreographies, it is allowed to arrange the connections freely.

Then they evaluated the difficulty, practicality, and memorability of the choreographies by five levels and reported the reasons. Difficulty is for considering the technical level of a dancer who can perform the sequence. Practicality is related to human body features, and how easy to make the movements. Ten sequences are created by automatic composition system. Two of ten sequences has arrangement of beat. Another two of ten sequences are constructed by contemporary movements only. The other six sequences are composed by blending ballet steps and contemporary dance movements without beat arrangement. Figure 4 shows an example of a created sequence by blending ballet steps and contemporary dance movements.

3.4. Results and discussion

Each of the three experimental subjects rated each of the ten sequences, yielding 30 ratings for each factor.

The average rating of difficulty was 3.6. This was varied among sequences. The rating contains the levels between "an advanced student could accomplish" and "a professional dancer could accomplish." No sequence was evaluated as "cannot be performed."

In the case of practically, the average rating of three dancers was 3.4. The rating contains the levels between "normal" and "easy to perform." The main rea-

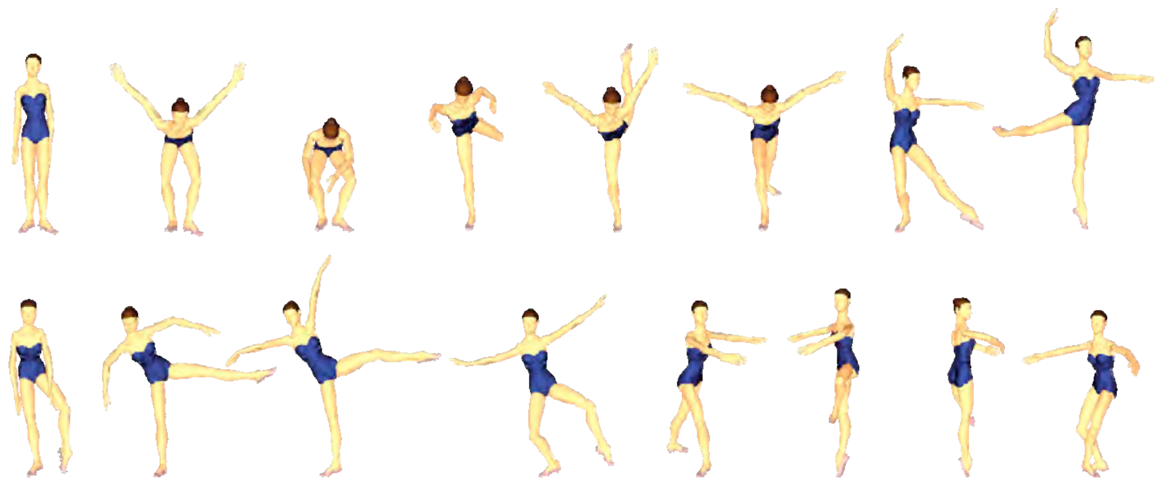


Figure 4. Example of created sequences by automatic composition

son why they felt difficult to perform was the occurrence of a ballet step between contemporary dance movements. In the case of contemporary dance only, no movement was evaluated as "difficult to perform."

In the case of memorability, the average rating of the three dancers was 3.4. The rating contains the levels between "normal" and "easy to memorize." The rating was different between dancers. The reasons why they felt "difficult to memorize" were the repetition of steps/movements with no specific features, and the continuing same steps/movements for a long time. The repetition of difficult steps/movements was easy to memorize.

The arrangement of beat and the blend of dance genres create unexpected choreographies. For example, a fast and big jump without preparation was unexpected. This choreography has the possibility to make the audiences surprise if he/she performs it on stage.

4. Conclusion

In this research, we tried an interactive composition and an automatic composition of dance sequences by blending ballet steps and contemporary dance movements. We concluded that our proposed system supporting "analytic-synthetic composition" is useful, in the way that it gives a choreographer some kind of seed of choreographs. The system affords the 'heuristic' method for discovering a novel movement.

Our motion archive includes only 58 movements for contemporary dance at the moment. None the less, the choreographer never comes up with some novel series of movements, unless using this system. After adding more movements, we believe that the system will become more useful for choreographers.

In future works, we aim to develop a system that can automatically create some sequences like a specific choreographer by introducing the function of the choreographer.

References

[1] K. Perlin et al. Improv: A System for Scripting Interactive Actors in Virtual Worlds. Proc. of SIGGRAPH 96. pp.205-216, 1996.

[2] K. Pullen et al.: Motion Capture Assisted Animation: Texturing and Synthesis. Proc. SIGGRAPH2002. pp.501-508, 2002.

[3] T. Kim, S. Park, S. Y. Shin: Rhythmic-Motion Synthesis Based on Motion-Beat Analysis. Proc. of SIGGRAPH 2003, pp.392-401, 2003.

[4] J. James, T. Ingalls, et al.: Movement-based Interactive Dance Performance. Proc. of the 14th annual ACM international conference on Multimedia, pp. 470-480, 2006.

[5] Credo Interactive. LifeForms. http://www.credo-interactive.com/

[6] L. Wilke, R. Ryman, T. Calvert, I. Fox. Applications of computers to dance. IEEE Computer Graphics and Applications. 25(2), pp.6-12, 2005

[7] A. Hutchinson: Labanotation, or, Kinetography Laban, 3rd rev.ed. Theatre Arts Books, New York, 1984.

[8] A. Soga, B. Umino, et al.: Automatic composition and simulation system for ballet sequences, The Visual Computer, Vol. 23, No. 5, pp. 309-316, 2007.

An Enactive Interface for Playing Physical Sound Synthesis of Bowed String Instruments

Tapio Takala
Helsinki University of Technology, Finland
E-mail: tapio.takala@tkk.fi

Abstract

The paper presents a sound synthesis model and controller for bowed string instruments. As playing a string instrument needs continuous action with the bow, it cannot be properly done with discrete actions such as with a keyboard. Here a new enactive controlling device built on a real violin is presented. It allows normal playing gestures (bow motion and fingering) to be used, and provides natural haptic feedback to the player. The synthesizer core is a simulated vibrating string driven by bow friction. Unlike most string models based on the waveguide synthesis, this model uses the finite element approach, consisting of a chain of masses and springs, enabling a very direct and intuitive control. A variation of the controller has also been built for a musical hand saw. Based on testing both subjectively and in public performances, the synthesizer and controller together make a playable musical instrument.

1. Introduction

With traditional notation, music is described as discrete events (notes) in symbolic form. In practice this needs interpretation. Digital encoding (MIDI) allows finer control of onset timing and volume, which may be adequate for plucked or percussive instruments such as piano, guitar or drums, but still lacks the continuous tuning needed for wind and string instruments or vocal voice. Although attempts have been made to describe these with additional notation [10], enactive control is needed for full expressivity.

This is why bowed strings have long been a tough challenge to digital music. Repeating a simple sawtooth waveform, idiosyncratic to the Helmholz motion of a bowed string, does not reproduce the timbral evolution of a sound from an initial attack to final decay. Traditional approaches (sampling and FM synthesis) do not offer enough flexibility for changing the timbre dynamically. Physical modeling, although a more natural approach, is hard to control because of a multitude of non-intuitive parameters.

In my opinion the main problem is lack of adequate controllers. The characteristic sound of a bowed instrument comes more from its typical use than from its detailed physical properties. As a violinist myself, I would like to play a synthesizer like a real violin. An attempt towards that was presented recently [9]. This paper reviews the development and extends it with further detail and variation to other bowed instruments.

2. Enactive controller

The violin bow allows sensitive gestural control for producing various bowing styles, such as *legato*, *detache*, *spiccato*, *staccato*, *martele*, etc. Using a keyboard to turn the bow excitation of a physical model on/off is unacceptable, but we need continuous control. In order to keep playing experience close to a real violin, I didn't want to change the instrument radically. Also in order to avoid disturbing mechanical attachments and additional weight in the bow, I decided to leave the bow as it is, and to put all sensory instrumentation into the violin body. The arrangement (figure 1) consists of two parts, a thin rod simulating each string at the place where it is bowed, and a real but non-vibrating string to sense left hand fingering.

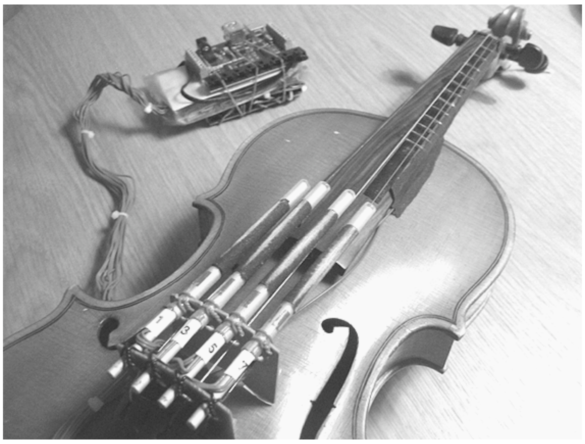


Figure 1. Prototype interface with bowing sensors for four strings and a single shared fingerboard sensor.

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

2.1 Related work

Many different mechanisms have been proposed as controllers for synthetic string instruments. Poepel and Overholt [6] present a review of recent developments.

Trueman and Cook [11] made a number of devices for sensing the bow and fingers utilizing e.g. accelerometers and force sensitive resistors (FSR). They also experiment with the sound reproduction through an instrument body containing loudspeakers. These developments lead to playing techniques and music styles clearly different from traditional violin.

Florens [2] used an electromechanical haptic device for bowing. It senses the hand motion, while also giving force feedback about the string's vibration. The response appears to be very natural, although moving in a smaller scale than real violin bow.

Special controllers aiming at traditional playing have been presented by Nichols [5] and Young [14]. Both utilize special sensors attached to the bow. Serafin *et al.* [7] suggest gestural mechanisms such as scratching with a graphical tablet to simulate bowing.

2.2 Sensing the bow motion

The two values essential for describing bow action are its pressure and velocity relative to the string. The dynamic proportions of these parameters (the bowing gesture) determine the necessary attack and release transitions idiosyncratic to each bowing style.

The bow controller presented here consists of an aluminum rod with two strain gages in orthogonal directions. One measures the vertical force caused by bow pressure, while the other measures the transversal force caused by the bow's friction when moving over the rod.

The device is not measuring exactly the bow velocity but the horizontal force based on friction and thus dependent on both pressure and velocity. This correlation is partly compensated in software in order to make these control quantities more independent.

The device works well for ordinary bowing styles with continuous motion (*legato*, *detache*, *spiccato*). However, those bowing styles where a note starts with the bow not moving but only pressing the string (*martele*, *staccato*) have the difficulty that an unintended horizontal force causes the model to play even if the bow is not moving. On a real instrument this force just pulls the string without making any sound. Solving this problem properly calls for further research .

2.3 Detecting the finger position

In order to determine the vibrating string length, which in turn defines the played pitch of a physical model, the position of left hand fingers is needed. Various options were considered, including force sensitive resistors (FSR) and resistive surface material on the fingerboard, but no suitable mechanism for continuous position control was found. Instead, the final solution is based on metallic frets attached to the fingerboard, connected in chain by fixed resistors. The metal coated string touching a fret makes a shortcut and yields a well-defined resistance to be measured.

The controller is easy to use and helps to play correct pitch easily. In addition, a real string as part of the controller has the advantage that it gives natural haptic touch to the fingers.

On the other hand, frets are not a natural arrangement for the violin, as they do not allow fine tuning of the pitch. Thus they prevent *glissando* effects, and force vibrato to be made artificially. The controller also does not measure the finger pressure, which would be needed to play *flageolettes*.

3. Physical modeling of string and bowing

Physical modeling of violin and other bowed strings has been studied a long time. An excellent review is given by Woodhouse and Galluzzo [13], referring to both important physical research before digital synthesis and to recent approaches, particularly concerning the bow-string interaction.

Serafin [8] concentrates on detailed models of friction between bow hairs and the violin string. Later she also has extended her research to frictional sounds other than those from the violin family.

An important issue is playability of a model. Guettler [3] measured it as the time interval in which the string starts to make Helmholz motion. This is mapped in the parameter coordinates of bow velocity vs. pressure into areas where most natural playing happens.

He also explains the physical phenomenon of subharmonics, used to bring special sounding effects by playing with very high pressure and slow movement of the bow.

3.1 The string model

Acoustic vibration of a string can be modeled as a chain of point masses connected by springs, which is a finite element approximation of the continuous Helmholz differential equation. Commonly it is formulated as a digital waveguide with appropriate filters. In this study, however, the mass-spring model is simulated directly, as is done in the CORDIS and GENESIS systems [1][2]. This provides more direct control and intuitive visualization of the string shape, making it particularly useful for interactive experimentation.

Considering transversal bending motion only (figure 2), the force acting on each mass (i) is the difference of spring tensions on each side of the mass, i.e. the force (F) is spring constant (k) times the transversal position (z) difference to each of its neighbors. Integrating the induced acceleration in time gives instantaneous velocity and position, which in turn determines the spring tensions. Normalization and discretization for unit masses and time steps results in the same simple formulas as used in waveguide models [12]. In addition, there is a damping force proportional (by constant ρ) to the velocity of the mass relative to its neighbors.

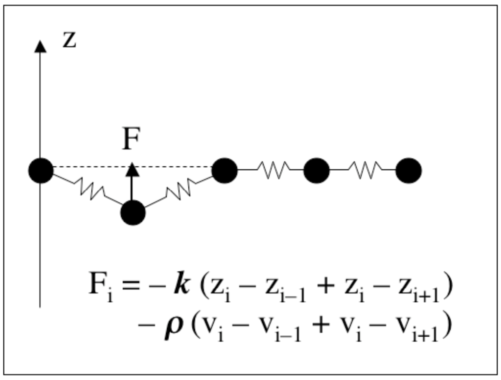


Figure 2. Mass-string model, with equation for the transversal force acting on mass *i*.

Touching the string with a finger is modeled as additional damping at an appropriate mass. A finger positioned between two masses damps each one with an interpolated factor. With a slight damping only at a proper nodal position, it is also possible to produce harmonic overtones (*flageolettes*).

The violin body is not modeled here at all. Although it filters the sound with formants, it is stationary and does not much contribute to the timbral evolution that makes violin sound distinguishable.

3.2 Bow friction

The modelling parameters most affecting the sound of the string are material and finger damping (determining the spectral decay rates, and that way brightness), violin body resonances (formants), and the coupling of string to the bridge (sympathetic vibration).

However, more important than string and body, is the bow and its interaction with the string. The control values needed are bow pressure and velocity, and (to a lesser extent) distance from the bridge. The relative values of pressure and velocity roughly determine what kind of sound is played (harsh for high pressure, whistling for high velocity), and there is an optimal relation range for best playability [13]. The dynamic shape and synchronization of these values (the gesture) determines a bowing style.

While the bow hairs move along a string, they alternatingly stick and slide, causing a typical sawtooth form of motion on the string. This frictional interaction is governed by a non-linear relation for bow pressure and the instantaneous velocity difference between bow and string [8], as depicted in figure 4.

The model has a large number of different parameters governing its behaviour. As the purpose here was not to make exact acoustical replicas of existing instruments or recorded sounds, the system was tuned experimentally. The goal was to make a sound audibly similar to real violin, based on my subjective assessment. The parameters are strongly interrelated, and much trial-and-error was needed. A complete analysis remains for future research.

An interesting curiosity that the model is able to reproduce is subharmonics, a string vibrating with lower frequency than its natural resonance. The phenomenon is explained by Guettler [3], but to my knowledge not synthesized before.

3.3 Interactive visualization

For understanding the model's behavior and for fine-tuning it, an interactive development environment was necessary. Part of it is graphical visualization of the string shape. Figure 3 shows a snapshot with four violin strings represented with point masses. Another visualization in figure 4 is the commonly used phase diagram relating the string's transversal velocity (horizontal axis) to the instantaneous force of the bow acting on the string (vertical axis).

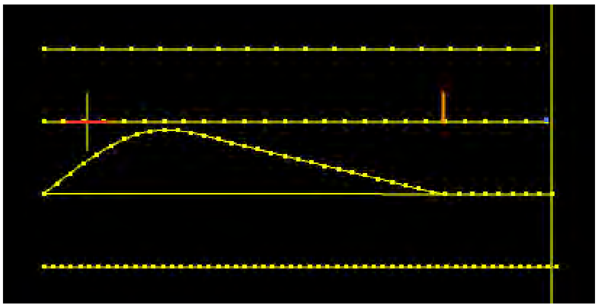


Figure 3. Snapshot from interactive simulation of the four strings of a violin (bow and finger positions shown by markers at left and right, respectively).

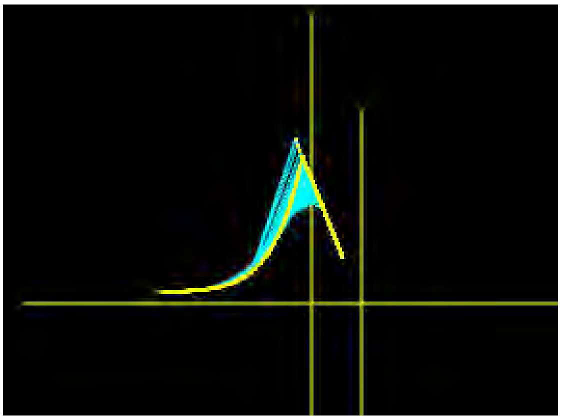


Figure 4. Real-time phase diagram displaying bow friction (x = velocity difference, y = bow pressure).

Although the mass-spring model is not optimal in view of computational efficiency, the current software running in an ordinary laptop is able to handle in real time up to 80 point masses, corresponding to two or three violin strings simultaneously.

4. A musical hand saw

Obvious extensions of the concept are bowed string instruments other than violin. Another variation is a musical hand saw attached with a similar bowing sensor but the synthesizer pitch controlled by a flexion sensor (strain gage) on the saw blade (figure 5). It has been played in a test performance, giving a believable musical effects, although a one-dimensional string does not appropriately model the vibrating blade.

Playing a correct pitch just by bending a blade requires a lot of practicing. In order to make this easier, a non-linear mapping from blade bending to musical pitch was applied, as was done by Kessous [4] for a tablet-based pitch controller.

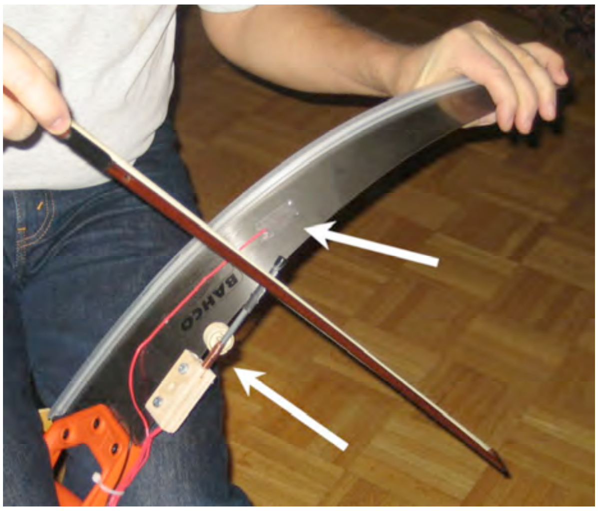


Figure 5. Digital hand saw in action (placement of bowing and bending sensors shown by arrows).

5. Conclusions

A sound model and physical controller have been presented for real time performance with bowed string instruments. The work combines previously developed modeling techniques with an enactive controller, and presents a usable instrument prototype. The various parameters governing a physical model's behaviour have been tuned experimentally, such that the model is playable within a wide range of controller input values. As a consequence, it enables the temporal variation necessary for lively musical performance with different bowing styles. In particular, the most common bowing styles (*legato*, *detache*, *spiccato*) perform well, whereas there are problems yet to be solved with those styles requiring a stop motion (*martele*, *staccato*).

The synthesis model and controller have been tested in practical use and found to be functional. The device has also been publicly demonstrated in a concert performance at RAFLOST Festival of Gesture Controlled Instruments, Reykjavik, in May 2007.

References

[1] Castagné, N, Cadoz, C. “GENESIS : A Friendly Musician-Oriented Environment for Mass-Interaction Physical Modeling”, *Proc. ICMC’02*, Göteborg, Sweden, 2002.

[2] Florens, J-L. “Expressive Bowing on a Virtual String Instrument”, in: Camurri and Volpe (eds.): *Gesture-Based Communication in Human-Computer Interaction* (GW 2003). LNAI 2915, Springer, 2004, pp. 487-496

[3] Guettler K. *The Bowed String – On the Development of Helmholtz Motion and On the Creation of Anomalous Low Frequencies*. Doctorate thesis, Royal Institute of Technology, Stockholm. 2002.

[4] Kessous L, "Bi-manual mapping experimentation, with angular fundamental frequency control and sound color navigation", Proc. NIME-02, Dublin, Ireland, 2002.

[5] Nichols C. *The Vbow, An Expressive Musical Controller Haptic Human-Computer Interface*. Doctoral thesis, Stanford University, 2003.

[6] Poepel C, Overholt D. “Recent Developments in Violin-related Digital Musical Instruments: Where Are We and Where Are We Going?”, *Proc. NIME’06*, Paris, 2006, pp.390-395.

[7] Serafin S, Burtner M, Nichols C, O’Modhrain S. “Expressive Controllers For Bowed String Physical Models”, *Proc. DAFx’01*, Limerick, 2001.

[8] Serafin S. *The Sound Of Friction: Real-Time Models, Playability And Musical Applications*. Doctoral thesis, Stanford University, 2004.

[9] Takala, T. "A playable synthesis model and controller for bowed string instruments", *Proc. ICMC'07*, Vol.1, Copenhagen, 2007, pp. 382-385.

[10] Takala T, Laurson M, Välimäki V, Hiipakka J. “An Expressive Synthesis Model for Bowed String Instruments”. *Proc. ICMC’2000*, Berlin, 2000, pp.70-73.

[11] Trueman D, Cook P. “BoSSA: The Deconstructed Violin Reconstructed”, *Proc. ICMC’99*, Beijing, 1999, pp.232-239.

[12] Välimäki, V and Takala, T: “Virtual Musical Instruments - Natural Sound Using Physical Models”, *Organised Sound*, Vol.1, No.2, Cambridge University Press 1996, pp.75-86.

[13] Woodhouse J, Galluzzo P. “The Bowed String As We Know It Today”. *Acta Acustica / Acustica*, Vol. 90, 2004, pp. 579-589.

[14] Young D. “The Hyperbow Controller: Real-Time Dynamics Measurement of Violin Performance”, *Proc. NIME’02*, Dublin, 2002.

Virtual Reality and Augmented Reality art explained in terms of sensory-motor coordination.

Michel F. van Dartel   Jan M. V. Misker   Anne M. M. Nigten   Jelle van der Ster
V2_ , Institute for the Unstable Media, The Netherlands
E-mail: michel@v2.nl, jan@v2.nl, anne@v2.nl, jelle@v2.nl

Abstract

Many contemporary electronic artworks deal with the creation of Virtual Reality (VR) or Augmented Reality (AR), but how viewers cognitively construct such realities is not commonly understood in the arts community. The latter is required to make an informed choice between VR and AR and decisions regarding how these realities should be realised. This paper will use VR artwork DEVMAP *(by Workspace Unlimited) and AR artwork* Exercise in Immersion 4 *(by Marnix de Nijs) to explain how viewers cognitively constructed VR and AR realities on the basis of sensory-motor coordination. Furthermore, it concludes by providing a rule of thumb for artists working with VR and AR; the strongest sense of reality is created by maximising the similarity in sensory-motor patterns that exist between interacting with the real world and interacting with the artwork.*

1. Introduction

Many contemporary electronic artworks deal with the creation of virtual and augmented realities. Most of these artworks aim to provide a viewer with an alternative reality that evokes a strong sense of reality, and, ultimately, a state of *immersion*. (By *immersion* we mean a state in which the user ceases to be aware of that he or she is acting in a different reality than the one in which his or her body is physically situated.) In the creation of such alternative realities, the choice between Virtual Reality (VR) and Augmented Reality (AR), and decisions regarding the realisation of such a reality, strongly determine the sense of reality that will be experienced by the artwork's viewers, and are thus crucial regarding the translation of an artistic concept into an artwork. To make an informed decision regarding the type of reality to use, and decisions regarding the realisation of the reality, it is useful for an artist to understand how viewers cognitively construct VR and AR realities. Therefore, this paper aims to bring theory to practice by informing practitioners in VR and AR art on how VR and AR realities are cognitively constructed.

Below, VR artwork *DEVMAP* by Workspace Unlimited (which was commissioned by V2_) and AR artwork *Exercise in Immersion 4* by Marnix de Nijs (which is currently in development at V2_ and part of the Dutch BSIK project MultimediaN[1]) will be used to explain how viewers cognitively constructed VR and AR realities on the basis of sensory-motor coordination. Although the paper will limit itself to discussing the visual aspects of VR and AR artworks, its conclusion may be generalised to other modalities.

To provide the required theoretical background, section 2 will discuss the sensory-motor coordination approach and how this approach explains the cognitive construction of reality. The following section, section 3, will briefly describe the difference between VR and AR and use *DEVMAP* and *Exercise in Immersion 4* to explain how viewers cognitively constructed VR and AR realities on the basis of sensory-motor coordination. Section 4 will then discuss when VR or AR may be preferred in art, and concludes by providing a rule of thumb for artist using VR or AR.

2. Construction of reality through sensory-motor coordination

Although already widely claimed in 1970s ecological psychology by J.J. Gibson [4], relatively recent research proves that our sense of reality is constructed through active interaction with the environment, rather than by passively receiving information from the environment. It has been shown that such active interaction enables storage of knowledge about the environment in the coordination between sensory and motor systems [3] and that such coordination enables conscious experience of the environment [7]. In other words, by coordinating our sensory and motor systems we construct the reality of the world around us.

To give a simple example. When we see an apple for the first time, we use our bodies (i.e., our motor systems) to explore its shape and surface (even if only

[1] See www.multimedian.nl for more info.

glancing our eyes over it). The pattern of changes that our senses receive due to this exploration (such as the changes in light on our eyes' retinas) will from now on be characteristic for an apple. Any new encounter with an apple will result in a more or less similar exploration, resulting in this characteristic (i.e., invariant) pattern of sensory changes. Together, this exploration and the resulting patterns of change now constitute our conscious experience of apples.

Conveniently, many things in our world have similar characteristics (such as apples and apple green sweaters for instance) and we can generalise stored sensory motor patterns associated with a certain characteristic (such as an apple's green-ness) to other things with that same characteristic, resulting in a similar conscious experience of that characteristic upon encountering it. The latter makes it relatively easy for us to store new objects, since almost all new objects in the real world have characteristics that we are already familiar with. This enables us to construct the reality of our everyday world in relative ease, and upon the rare occasion that we do encounter something completely new, we can use our existing sensory-motor coordination to base our new experience on. Hence, the fact that humans often express the thought that this new such-and-such looks, sounds, or feels very much like this-and-that.

In essence, this is how we construct our reality, whether this reality is real, virtual, or augmented.

3. Virtual Reality and Augmented Reality

The difference between VR and AR is that in VR the whole environment is simulated [1], while in AR the real surroundings are only partially overlaid with simulated elements [5].

In order to provide insight into the sense of reality brought about by VR and AR artworks, this section will briefly describe how virtual and augmented realities are constructed in terms of sensory-motor coordination. We will do so on the basis of two concrete examples of art projects that deal with VR and AR: *DEVMAP* by Workspace Unlimited and *Exercise in Immersion 4* by Marnix de Nijs, respectively. Although these artworks have very different conceptual aims (that this paper will not go into) both aim to convey their concepts by providing an alternative reality that evokes a strong sense of reality, and ultimately, a state of immersion.

3.1 Virtual Reality: *DEVMAP*

The VR installation of *DEVMAP* and its users' interaction with the installation is typical for many VR artworks. A user stands in front of a screen that is placed a few meters away from him or her. Users control the movement (walking) of an avatar in the projected virtual world by moving a mouse (walking direction) and

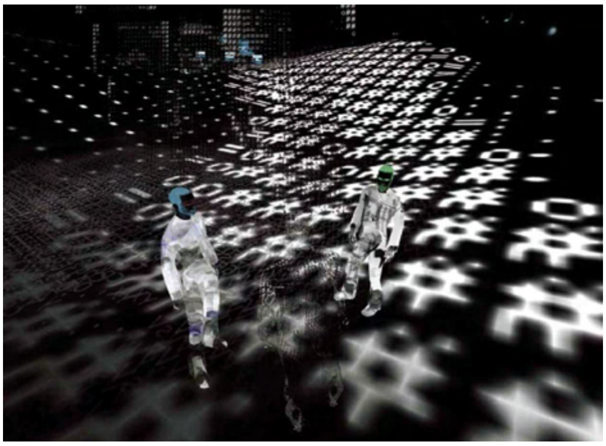


Figure 1: Screenshot from *DEVMAP* at the Dutch Electronic Art festival 2004

by pressing the left or right mouse buttons (forward or backward walking, respectively). The structure of the *DEVMAP* virtual world is made up of data (audio, video, or text) which can be streamed live into the virtual world (as was done during the Dutch Electronic Art Festival[2] 2004, see figure 1).

Since humans construct their realities through active interaction with their environment and base new experiences on sensory-motor patterns that they are already familiar with (see section 2), the (initial) sense of reality experienced in a VR or AR artwork is determined by the similarity in sensory-motor patterns that exist between interacting with the real world and interacting with the artwork. The smaller this similarity is, the harder it will be for a user to generalise sensory-motor patterns that he or she is familiar with from the real world to the other reality, and the weaker the experienced sense of reality in the VR or AR world will be.

Since in VR the whole environment is simulated (rather than only partially overlaid with simulated elements as is the case with AR), most sensory-motor patterns will be new to a first-time user. Therefore, the sense of reality in VR artworks is often not very strong at first, but grows with a user's interaction with the simulated environment. Such is also the case with *DEVMAP*. Upon a user's first interaction with the *DEVMAP* virtual world, navigating through the *DEVMAP* environment feels more like scrolling through a text document than like walking through a landscape. However, the sensation of reality experienced while interacting with the *DEVMAP* virtual world grows on a user after a period of actively engaging with the VR environment, as the new sensory-motor relations become more familiar. This growing familiarity with the sensory-motor rela-

[2]The Dutch Electronic Art Festival is a biennial international and interdisciplinary festival organised by V2_, Institute for the Unstable Media.

tions of interacting with the VR environment can be considered a growing *suspension of disbelief* [2], as a user's willingness to overlook the abstract nature of the *DEVMAP* virtual world grows. Despite this growing sense of reality, after a long period of interacting with the *DEVMAP* virtual world, the experienced sensation is still not very comparable to the sense of reality felt in the real world, and a state of immersion is still very far off.

This can be attributed to the fact that only a few sensory-motor relations can be generalised from the real world to the *DEVMAP* virtual world. When we move our physical bodies in the real world, the changes in stimuli that are received due to that movement are invariant. As illustrated in section 2, our perceptual system uses these invariants to make sense of the world (i.e., they enable experience through sensory-motor coordination). When we move our avatar in the *DEVMAP* virtual world the changes in stimuli that are received due to that movement are completely different from those that would have been received in the real world. Thus, invariants that the user is familiar with cannot be used. Besides that, the user is also confronted with the fact that he or she interacts with two non-overlapping realities at the same time, manipulating a mouse in the real world and controlling an avatar in the virtual world. This means that the user has to ignore the real world sensations in order to feel a sense of reality in the virtual world.

In this line of reasoning, the sense of reality in *DEVMAP* would greatly benefit from a projection that changes according to a user's physical movement (precisely as it would change if that movement was made in the real world) and that shuts off any sensory feedback from the real world. Therefor, many VR installation use a Head-Mounted Display (HMD) or a Cave Automatic Virtual Environment (CAVE) to project images according to one's bodily movement.

Although real-world images would work best to recreate real-world invariants, the technology to reconstruct 3D models from real-world images is still very much in its early stages, and can be considered too complicated (for even the smallest space under the most basic light and surface conditions) and expensive (as such models are mostly created manually) for most art projects.

Despite the above, a fair sense of reality can be created using VR by allowing as many generalisations of existing sensory-motor relations as possible, and by persuading a user to intensively interact with the virtual world over a long period.

In a sense, entering a VR for the first time is very much like a baby entering the real world and being confronted with all new sensory-motor patterns. Both the baby and the first-time user start by observing the consequences of explorative movement (or reflexes), then gradually learn to coordinate their motor and sensory system to make sense of the world. Since it takes weeks or even months for a baby to pass this *reflex schema stage* [8] one can imagine how long it takes a first-time VR user to obtain a strong sense of reality in a VR world. However, as a baby may only be able to in later stages of its cognitive development, a first-time user can base new relations between his/her movement and the resulting sensory feedback of the VR world on the sensory-motor relations that the user is already familiar with.

In this light, in order to achieve the strongest sense of reality, it makes sense to mimic as many real-world sensory-motor relations as possible in the alternative reality to allow users to make use of the invariants that they are already familiar with. This explains why AR artworks are, in general, capable of providing a much stronger sense of reality than VR artworks. AR exploits the use of real world invariants, not only by mimicking sensory-motor relations in the overlay of virtual elements, but by retaining most of these relations by making the real surroundings part of the alternative reality.

3.2 Augmented Reality: *Exercise in Immersion 4*

The AR of *Exercise in Immersion 4* (*EI4*) is a spectacular artistic endeavor into partially overlaying real surroundings with simulated elements. The *EI4* art-game was first demonstrated as a prototype at the Dutch Electronic Art Festival 2007 exhibition, where a deserted storage building was used as the real surrounding to overlay with simulated elements (see figure 2). *EI4*

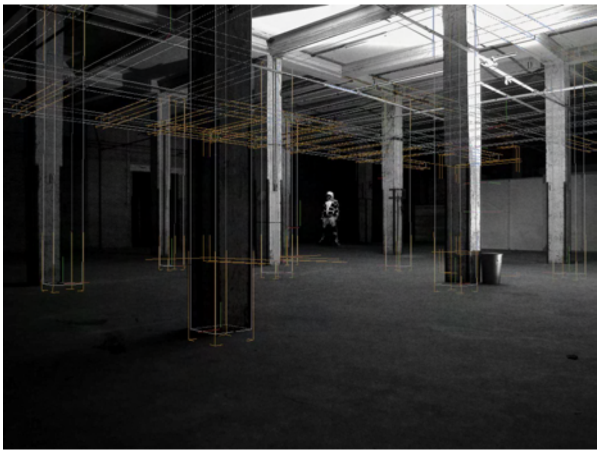


Figure 2: Screenshot from *Exercise in Immersion 4* at the Dutch Electronic Art Festival 2007

players wear a specially designed headset-display and a crash-suit. The headset has a sensorsystem that connects the position of a player with previously modeled visuals. A player starts the game in common reality, but the common reality is increasingly visually taken over by simulated elements as he/she progresses in the game.

The goal of the game is to collect *bionts*, small virtual balls that float around in the air. Collected bionts gather in front of the player, and bounce off real and virtual obstacles. When enough bionts are collected, the player progresses a level in the game. With every subsequent level, the player is more exposed to a virtual world. At later stages in the game, the virtual world even completely occludes the real world and real-world obstacles are not visible to the player anymore. Then, the bionts serve as a navigational aid, avoiding crashing into obstacles like walls and pillars in the real world.

As was already stated in subsection 3.1, the main advantage of AR over VR is that most of the invariants that users are familiar with in the real world are retained, with only new sensory-motor patterns to learn for the overlaying simulated elements. Therefore, as long as the surroundings are only partially overlayed with simulated elements, the sense of reality in AR artworks is by definition stronger than that of VR artworks. However, the main problem with AR is that the sense of reality in AR artworks strongly depends on the level of synchronization of the simulated elements with the real world. If the overlay of the simulated elements runs asynchronous with the real world, then the invariants that the user relies on to make sense of the reality become variant, and the sense of reality for the simulated elements is lost. Unfortunately, synchronizing the simulated elements with the real world also poses the main technical challenge of AR, since in order to do so the movement of a user needs to be very precisely tracked.

During the development of *EI4* the accuracy in tracking was also one of the main technical obstacles of the project. At the time of testing the *EI4* installation at the Dutch Electronic Art Festival 2007, the accuracy was good enough to achieve an AR experience with a strong coherence between the real surroundings and the simulated elements, resulting in a fairly strong sense of reality and, according to some users, a state of immersion.

The first tests at the festival also made it clear that the mismatches and glitches between the real surroundings and the simulated elements that were still present in the installation at the time of testing were, somewhat surprisingly, not disturbing the sense of reality very much. This *suspension of disbelief* [2] can, on a perceptual level, be explained by that our conscious experience of perceiving everything in our world in detail is in fact an illusion based on the *immediate availability* of everything in detail in the real world. As long as we do not direct our attention toward certain details, we will not be aware of these details but still experience the scene as a whole, even if things are missing or not completely right [7]. The developers were, for obvious reasons, directing their attention to mismatches and glitches in the simulated overlay. In contrast, users treated the AR world as if it was the real world, moreover not noticing the disturbances in the invariants used to construct their reality. This underlines that AR's use of real-world sensory-motor relations creates a stronger sense of reality, even preventing disturbances in the invariants from disturbing our sense of reality.

4. Discussion and conclusion

In the above we explained how the sense of reality in VR and AR artworks is determined by the similarity in sensory-motor patterns that exist between interacting with the real world and interacting with the VR or AR artwork. Since this similarity is larger for AR than for VR, it may be argued that if the aim of the artwork is to provide a realistic alternative reality, then AR may be preferred over VR. However, if the project does not aim to do so, or the main aim of the artwork is to provide a completely new reality (and the realistic nature of the artwork is only secondary), then VR may be preferred over AR. In the latter case, the similarity in sensory-motor patterns that exist between interacting with the real world and interacting with the VR artwork may be enhanced by the use of tracking systems and 3D projection techniques, such as HMD's and CAVE's.

We conclude that the rule of thumb to use for artist is that the strongest sense of reality is created by maximising the similarity in sensory-motor patterns that exist between interacting with the real world and interacting with the artwork.

References

[1] Burdea, G., Coffet, P. Virtual Reality Technology. Wiley-IEEE Press, New York, NY, 2003.

[2] Coleridge, S. T. Biographia literaria. 1817.

[3] van Dartel, M. F. Situated Representation. Universitaire Pers, Maastricht, 2005.

[4] Gibson, J. J. The Ecological Approach to Visual Perception. Houghton-Mifflin, Boston, MA, 1979.

[5] Haller, M., Billinghurst, M., Thomas, B. Emerging Technologies of Augmented Reality: Interfaces and Design. Idea Group Publishing, Hershey, PA, 2006.

[6] Maturana, H. R., Varela, F. J. The Tree of Knowledge - The Biological Roots of Human Understanding. Shambhala, Boston, MA, 1987.

[7] O'Regan, J. K., Noë, A. "A sensorimotor account of vision and visual consciousness.", Behavioral Brain Sciences 24, pp. 939 - 1031, Cambridge University Press, Cambridge, MA, 2001.

[8] Piaget, J. The Child's Construction of Reality. Routledge and Kegan Paul, London, UK, 1955.

[9] Ziemke, T., "The Construction of 'Reality' in the Robot." Foundations of Science, 6(1), pp. 163-233, Springer, The Netherlands, 2000.

One World: New Artistic Strategies for Integrating Interactive 3D Computer Graphics and Physical Space

Fabian Winkler* Laura Arns*† David Braun‡
(*)Department of Visual and Performing Arts, Purdue University, USA
(†)Envision Center for Data Perceptualization, Purdue University, USA*
(‡)Rosen Center for Advanced Computing, Purdue University, USA
E-mail:fwinkler@purdue.edu, larns@purdue.edu, dbraun@purdue.edu

Abstract

The dilemma of VR as medium for new media art has been that of its immaterial nature and an almost conscious denial of the material world, including interactors' bodies. This paper introduces a conceptual infrastructure, examples of artistic explorations and a concrete technological system for a strong integration of 3D computer graphics and physical space. It explains the three interrelated parts of our infrastructure: image, interface and physical space related to art and design practice. Relevant concepts of navigation, site specific simulation and reality jamming are illustrated in our actual artistic explorations and the paper closes with ideas for future project work based on a novel technological system interfacing the Arduino I/O board to VRJuggler and OpenGL, which can be of benefit to research and experimentation in both, the arts and technology.

1. Introduction / Context

Virtual Reality as art medium seems to have been a fad of the 1990s. The first public presentation of the CAVE [1] at SIGGRAPH 1992 and easier access to head mounted displays spawned a multitude of immersive VR artworks which all tried to seduce the audiences with new synthesized realities. Why did this trend not continue or even increase in popularity in the world of new media art? The technical infrastructure (hardware and software) has certainly improved over the past 15 years but the content on tiled wall displays, HMDs and CAVE screens has been largely reclaimed by scientific and military simulations and CAD models for engineering projects. In this paper we argue that one of the reasons for this development can be found in the way traditional computer generated immersive spaces have dealt with the interactor's body and his/her perception of the surrounding physical space: both have largely been neglected or even deliberately denied. Many VR worlds were designed to be experienced as (virtual/artificial) realities separate from the material world. The vision/brain apparatus was consciously disconnected from the rest of the body, which lead to a monopolization of the sense of sight (sometimes enhanced by sound) in the reception or experience of the artwork. In addition to unpleasant side effects, such a simulator sickness, this approach certainly missed many opportunities of exploring the expressive potential of a variety of forms of sensorial immersion and interaction. In his 1998 book „Virtual Realism“ [2], philosopher and designer Michael Heim has already argued that audiences can only then experience VR worlds more fully when they ground themselves more deeply in the physical world. In this paper we would like to extend this idea and propose a framework for an art practice that successfully integrates interactive 3D computer graphics and physical space. In the following sections we will introduce the infrastructure, technologies and resulting artworks that are based on our notion of a polysensorial environment, in which the virtual informs the real and vice versa.

2. Infrastructure

2.1.Image

The images in the works described below are realtime responsive, i.e. they mirror the behavior of the audience and in turn influence it. They are projected into the physical space and relate to it conceptually and aesthetically. Furthermore it is important to comment on the three-dimensional representation of the images' content: early realtime responsive images in the art, such as those created by Thomas Wilfred's Lumia in the 1930s followed the rules of 2D composition on a flat surface as developed in the long history of painting. The three-dimensionality of later computer generated images (see for example early examples in the field of media art such as: Michael Saup's sound responsive 3D forms in *pulse8*, 1992/94 or Knowbotic

Proceedings of ENACTIVE/07
4th International Conference on Enactive Interfaces
Grenoble, France, November 19th-22nd, 2007

Research's data space *Dialogue with the Knowbotic South*, 1994) allow to focus more on sculptural/architectural principles, building and navigation processes and "Mehransichtigkeit" (a form of spatial narrative unfolding while circumnavigating a sculpture). Interestingly enough, physical space is inherently three-dimensional and the computer-generated graphics should be considered as more than a framed window to the virtual.

2.2. Interface

The images can be manipulated by an interactor through a physical interface, which provides another link between the projected images and the physical space around them. In the concept and design of these interfaces we are often influenced by product design strategies of re-use and re-interpretation and appropriation, allowing the interactor immediate, tactile and often intuitive access to the functions of the interface. This approach is similar to the philosophy of Dutch design label "Droog": "Our products ... deal with memories, ..., re-use, craftsmanship, nature. They generate experience, interaction, participation, products that are easily to comprehend, have meaning, tell stories." [3] It also overlaps with some of the concepts of Japanese "Device Art" [4], integrating art and technology as well as design and popular culture. For example, the navigation interface in the work *architectural JoxMox – musical Encounters* (Fabian Winkler in collaboration with Jan Sturm, 2001) was based on the experience of skateboarding. Shifting the body weight on an elastically supported footpad allowed the interactor to navigate through computer generated three-dimensional structures.

Figure 1. Footpad interface, *architectural JoxMox – musical Encounters*. Photo credit: Jan Sturm

2.3. Physical Space

Frank Popper defines an artistically modified space as "environment": "a plastic statement contained within a space which possesses its own plastic 'meaning'." [5] This means that the physical space in which 3D computer graphics are placed already possesses attached meanings or conceptual properties, which interact with the meaning and conceptual properties of the interface and the computer generated image. The screen is where projected image and space intersect, it is more than just a planar projection surface with no connection or conceptual embedding in the exhibition space.

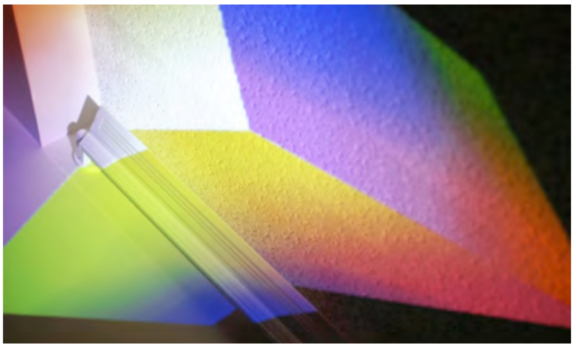


Figure 2. *amTRIance* (Michael Muehlhausen, 2007), *BE_TWEEN* exhibition Purdue University. Photo credit: Michael Muehlhausen.

Popper further divides "environments" in five categories: static/purely visual; kinetic, i.e. related to the activity of the spectator/visitor; polysensorial, with tactile and sonorous elements; natural (i.e. land art) and total environment, in which the visitor is witness of its behavior. Clearly, the research and analysis of kinetic and polysensorial environments are most relevant to the development of our future artworks interfacing 3D computer graphics and the material world.

3. Artistic Explorations and Concepts

How can interactive 3D graphics and physical space be integrated in meaningful and conceptually rich ways? The following examples briefly illustrate some approaches in Fabian Winkler's own work and as a result of the course "Envision Art 01: the responsive screen" (http://web.ics.purdue.edu/~fwinkler/590E) taught by the authors at Purdue University in the Spring 2007. This course, offered through the Electronic and Time-Based Art Program in Purdue's Department of Visual and Performing Arts, combined resources and technologies in the Envision Center for Data Perceptualization with contemporary artist practice and research.

3.1. Navigation

What does it mean to change computer generated structures and spaces by navigating in physical space? This was one of the questions Fabian Winkler's work *Cavallerizza Digitale* (2000) raised. The bodily experience of walking in the exhibition space (the baroque Cavallerizza Reale in Torino, Italy) was translated into deconstruction processes in a computer-generated architecture. The audience's navigation of the physical space became the interface between real and computer-generated architecture.

Figure 3. *Cavallerizza Digitale* (Fabian Winkler, 2000). Photo credit: Fabian Winkler

3.2. Site-specific simulation

The same work placed a responsive computer generated model of the interior façade into its physical counterpart. The movements of the audience in front of the real façade morphed individual architectural elements of the computer-generated façade into their geometric counterparts: ornamental baroque details were transformed into abstract bounding boxes. This process enhanced the audience's perception of both the physical architecture and the 3D model.

Figure 4. *Cavallerizza Digitale* (Fabian Winkler, 2000). Photo credit: Fabian Winkler

3.3. Spaces of encounter

Fabian Winkler's work *architectural JoxMox – musical Encounters* introduced another concept for mixing the boundaries between physical and computer-generated space. The computer generated 3D structures were used as interfaces, while navigating the virtual space (on the above mentioned footpad, see 2.2.), interactors could trigger sounds by colliding with the geometries. Other interactors controlled the shape and extension of the geometries through a physical interface (midi controller with turn knobs). The VR environment became a space for musical collaboration in the process of creating and navigating it.

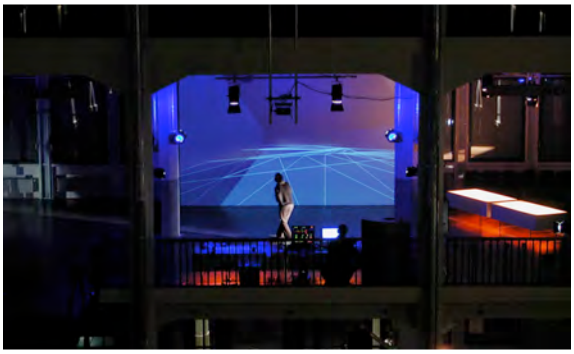


Figure 5. *architectural joxMox – musical Encounters*, ZKM Media Museum, Karlsruhe, Germany, 2001. Photo credit: Jan Sturm

3.4. Reality jamming

"Reality jam" is term we borrowed from one of the ISEA 2008 conference themes. It describes a "pressing together of the real and virtual in a context where their distinctions are deliberately obscured." [6] Students in our course "Envision Art 01: the responsive screen" dealt with this question in their works in surprisingly new ways, two of which we will briefly introduce:

3.4.1. Marbleometry

Jack Moreland's *Marbleometry* uses marbles as an input device to create geometries. The frequency and speed of marbles on a sensing marble track triggers drawing, extrusion and distortion sequences on a screen. The clicking sounds of marbles falling through holes in the marble track audibly accompanies the geometry shaping processes.

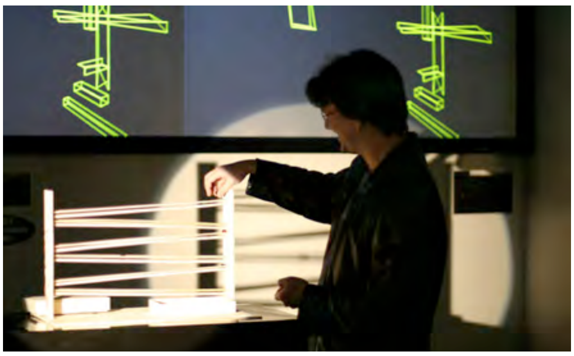


Figure 6. *Marbleometry*, *BE_TWEEN* exhibition Purdue University. Photo credit: Jack Moreland.

3.4.2. [re]CONNECT

Carlene Takaki's work [re]CONNECT is based on her observation that „we are always, in some way, reacting and adjusting ourselves to our surroundings, which does not only include our static environment but most likely also other people.“ [7] The project visualizes these connections by abstract lines that connect dots on a projection screen. Crossing laser beams on the floor in front of the screen enables interactors to add or remove connecting lines in the abstract three-dimensional composition.

4. Technology

4.1. MIDI/NCSA portfolio/Java3D

Fabian Winkler’s 2000/01 works were based on a system that interfaced MIDI data (generated from MIDI enable music instruments, STEIM’s BigEyes tracking system and other physical sensors) with custom Java3D applications. The responsive 3D geometries were created in 3D Studio Max and imported into the SceneGraph through NCSA portfolio, which also handled the MIDI data import.

4.2. Arduino/VRJuggler/OpenGL

For “Envison Art 01: the responsive screen,” students were encouraged to construct works that could be integrated into environments such as the CAVE. The VRJuggler [8] toolkit was already commonly in use at the Envision Center, thus we decided to build on it for the course. VRJuggler works with a variety of graphics toolkits such as OpenGL (www.opengl.org), and a variety of hardware such as CAVEs and head mounted displays. The open source Arduino development board (www.arduino.cc) presented an excellent platform for students to create their interaction devices. The technological innovation of our approach was to design a software interface that linked the Arduino development board and VRJuggler (consisting of both, a VRJuggler serial driver and an Arduino driver). A simple block structure protocol was implemented such that there are two basic messages: a configuration message used to configure the I/O parts on the development board, and a data message that contains data values that are to be set. The Arduino was then setup to send and receive data messages. The VRJuggler driver implements a listening thread that receives incoming data and parses the data messages into VRJuggler data types, such as digital and analog inputs. As sensors are read by the development board, the data items are used to manipulate graphics functions in OpenGL. At the beginning of the project we intended to also send digital outputs to the development board but VRJuggler currently does not support this functionality.

The technological system we have developed opens new possibilities in research and experimentation in both, the arts and technology. It can help engineers to create inexpensive yet effective prototypes of their project ideas. At the same time it allows artists to create highly individualized interactive works for environments such as CAVEs or large displays without being limited by a lack of technical expertise.

5. Conclusion/Outlook

Our further research and artistic explorations integrating interactive 3D graphics and physical spaces will be based on the technologies described in 4.2. since they offer a flexible and scalable system for connecting the physical world to immersive VR spaces. Ideally this system will be developed to also inlcude a means for events in the VR space to manifest themselves in the material world (e.g. through luminous, sonorous, kinetic, tactile, etc... responses). Conceptually we would like to further explore the ideas of navigation and construction/deconstruction in VR spaces. Our approach continues to be polysensorial – dealing with the human sensorium as an “ever-shifting social and historical construct” [9] that can be best explored in the realm artistic experimentation - through physical interfaces based on strategies of re-use, appropriation and re-interpretation and environments in which interactive 3D graphics are embedded into physical spaces in meaningful ways.

References

[8] Bierbaum, A., Just, C., Hartling, P., Meinert, K., Baker, A., and Cruz-Neira, C. “VR Juggler: A Virtual Platform for Virtual Reality Application Development”, Proceedings of IEEE VR 2001, Yokohama, Japan, 2001.

[9] Bull, M., Gilroy, P., Howes, D., and Kahn, D. “Introducing Sensory Studies”, The Senses and Society, vol. 1, no. 1, pp. 5-7, Berg Publishers, Oxford, UK, March 2006.

[1] Cruz-Neira, C. “Projection-based virtual reality: The CAVE and its applications to computational science”. Ph.D. Thesis, University of Illinois, Chicago, IL, 1995.

[3] Droog “Welcome to Droog”, http://www.droogdesign.nl (last checked: Oct. 10th, 2007).

[2] Heim, M. Virtual Realism, Oxford University Press, New York, 1998.

[6] ISEA2008 “Reality Jam”, http://www.isea2008.org/themes5.html (last checked: Oct. 10th, 2007).

[4] Kusahara, M. “Device Art: A New Form of Media Art from a Japanese Perspective”, intelligent agent vol. 6 no. 2 , http://www.intelligentagent.com/archive/Vol6_No2_pacific_rim_kusahara.htm (last checked: Oct. 10th, 2007).

[5] Popper, F. Art: Action and Participation, New York University Press, New York, 1975.

[7] Takaki, C. “[re]CONNECT”, artist statement, 2007.