Kungliga Tekniska högskolan

DT2140 mintint18 HT18-1 Multimodal Interaction and Interfaces

Do you trust your senses ?

Daniel DEL CASTILLO <danieldc@kth.se>
Camille FOURNIER<camillef@kth.se>
Tallulah GILLIARD <tallulah@kth.se>
Hemanth RAM KARTIK BOMMAKANTI <hrkbo@kth.se>

(group 33) KTH university

Referent teacher : Eva-Lotta

DT2140 mintint18 HT18-1 Multimodal Interaction and Interfaces

January 2019

ABSTRACT

This project focused on investigating how Augmented Reality could be a tool to experience how humans process uncorrelated information from their senses, particularly visual and haptic senses, and causes for confusion.
A real object with QR codes was placed inside a box. A hole on the side of the box allowed the subject to touch and manipulate the real object. Another hole on top allowed a phone's camera to see the contents of the box. When the camera identified the QR code, a virtual object, that could be the real one or a completely different object, was displayed on the phone with it's movements synchronized to that of the real object. Under no circumstances was the subject allowed to see directly inside the box.
Three subjects were tested with this setup and tried to guess the object inside the box on two different trials with different objects and answer some qualitative questions about their experience afterwards. The level of confusion was surprisingly high as they made wrong guesses half the time, although they mostly relied more on their sense of touch than their sight.

Keywords

AR, VR, touch, illusion, Unity, Vuforia

Number of words : 3275

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Touch and Illusions

The project focused on fooling the mind. The objective was to make the user experience confusion by having uncorrelated haptics and visuals, hence the work was based on papers about touch, haptics and virtual illusions.
First, some papers that were previously known of were investigated, focused on topics about mind and body illusions.

Ramachandran et al. [1] continued the line of investigation which was started by Ramachandran himself, based on the hypothesis that there is a link between the phantom limb phenomenon and human adult brain plasticity. In their work, they take this link as an opportunity to explore the topographical changes that occur in the cortices of amputee individuals by performing on them imaging techniques such as fMRI. The most relevant aspect of this work is given by the insightful experiments they designed and conducted, where phantom limbs are induced in non-amputee subjects, therefore demonstrating how the human brain can be fooled by feeding it a misleading combination of sensory information sources.

Henrik Ehrsson, in this famous 2008 paper [2] has done a series of 5 experiments using VR headset to display images from another point of view than the subject point of view. The images seen by the subject are the one from a mannequin point of view which the stomach is being touched at the same time of the subject stomach to create the illusion. Other objects can be used to be a threat to the subject (like a knife they get closer to the abdomen of the subject). They also tried if that the body would need to look like a human to be experienced as one's own by using tables or chairs which did not work. They have tried to see that people should be able to swap bodies with each other and, quite literally, to shake hands with themselves while experiencing ownership of another person's body. For each experiment they measure the results with skin conductance response and questionnaires. The results of our experiments demonstrate that healthy volunteers can perceive another person's body, or an artificial humanoid body, to be their own. This speaks directly to the classical question of the relationship between human consciousness and the embodied cognition (Varela, 1991 "The Embodied Mind"), which has been discussed by philosophers, psychologists, and theologians for centuries.

In [3], Matthew Botvinick and Jonathan Cohen conducted two experiments on perceptual processes when visual and touch input give different information. In the first experiment, the left arms of the subject was hidden to them and a rubber hand was placed in front of them. They tried to stroke both hands with a paintbrush at the same time to build the illusion the rubber hand was their own. It worked perfectly : the subjects did think of the rubber hand as their own and felt the paintbrush in the location the rubber hand was touched, and not where they were actually touched. the second experiment allowed them to know more about this distortion of position sense. They reproduced the same experiment, but asked their subject to align their right index with their left index by moving it in a straight line, both before and after the paintbrushes. A control group also did the experiment but the strokes were slightly delayed between the rubber hand and

their hand during the experiment. As a result, the experimental group displayed more reach displacement than the control group (23 mm toward the rubber hand compared to 13 mm away from it for the control group).
This paper show the possible dominance of visual input over touch input and also the importance of timing during timing during the experiment to make the mind accept visual input instead of touch input.

In [4], the authors propose a general principle for a brain mechanism that is central to our project: the discrimination performed during object recognition when sensory data from different modalities is adversary.
They analyze this mechanism for the case of visual and haptic competition, and they propose a maximum-likelihood estimator as the method for reproducing the experimental data with the lowest variance.
For the experiment, they collect the perceptual experiences from several users that have to estimate the size of two bars with different lengths, both through a visual and a haptic interface. Finally, modality dominance is given by the estimation with the lowest variance associated.
While the brain mechanism they try to explain is certainly relevant for this project, we will not apply the methodology described in this paper.

In [5], the authors demonstrate visual dominance over haptic for the particular case of mechanical stiffness perception. Similarly to our setup, they provide a virtual environment as the visual output. They do this by displaying virtual images of springs that deform in response to the users' forces applied on a haptic interface device. The stiffness of the haptic interface that reproduces the the mechanical behavior of the springs is controlled through a computer.
Later, the users are asked to judge the stiffness of each virtual spring they interacted with. The results show significant erroneous judgements when visual deformation of the softer spring was equal to the deformation of the harder spring for that force and vice versa. This is relevant to our project, since it indicates that an important aspect of haptic perception such as the stiffness of an object can be altered through controlled visual perception.

Fairhurst et al. [6] rightfully point out that it is human nature to have an inclination to touch what is seen as confirmation of their existence/reality. Based on the results from the tests conducted, it is that the users had a greater level of confidence in their decision when they touched the object rather than when they saw it (confidence rating). However, more accurate results were achieved from only seeing the object to identify its reality in terms of size and shape (perceptual sensitivity). The tests are significantly different from the one proposed in this project, but it would be interesting to see if similar characteristics are exhibited from people when they play the game that is created. The psychological aspects are also considered in this paper, which we will not be exploring for this project.

Lederman and Klatzky [9] have defined six manual "exploratory procedures" and their associated object properties. Since our physical tangible object will be hidden, our subjects will probably use theses procedures to explore and have information about what they are touching, grabbing or holding objects.

The group also checked how the dominance of visuals of haptic cues were processed by subjects. The Colavita visual dominance effect [12] is when participants respond more likely to the visual component of an audiovisual stimulus. Hecht and Reiner [13] found a similar visual dominance effect for haptic–visual stimuli.

1.2 Hypotheses

The aim of this research is to perform an experiment, where the mind would be confused by the virtual illusion uncorrelated with that of the physical touch.
So the research question that the group will attempt to answer here is if the user will be confused when visuals and haptics which are uncorrelated (what the user sees is different from what they touch)
This lead to the following hypotheses :
H1 : The participant will be confused about the object nature when uncorrelated haptics and visual cues are displayed.
H2 : The user will find easily guess the object nature when visuals and haptics are correlated.

2. MATERIAL & METHOD

2.1 Population

The subjects were recruited on KTH main campus. In total, 3 students (all males, age 18-25 years) accepted to be part of the experiment.

2.2 Material

Stimuli were presented through two mediums :

- A physical tangible real object (fruit or ball) hidden inside a box.
- A visual virtual environment displaying a similar object (fruit or ball) on a phone screen put on the box.

Hardware
The box is a cardboard box which dimensions are 30*32*19 cm.
An Android Phone was used to display the virtual environment, and a phone's flashlight along with LED lights to illuminate the inside of the box.
The real objects that were used included: a small basketball, a tennis ball, an orange, an apple, a kiwi and a grapefruit. They were recognized by the camera with 2 QR codes stuck on them in opposite sides.
The virtual objects displayed on the phone were: a basketball, an orange, an apple and a kiwi.
Software
The group developed and created an android application with the help of Vuforia v8.0.10 run on Unity 2018.2.15.

2.3 Experimental Design

To further disturb the subjects, some real objects were assigned different virtual ones on each sides. Each virtual object had a texture similar to the real object it was assigned to. Hence :

- The apple showed an apple or an orange

- The grapefruit showed an orange or a basketball
- The orange showed an orange or a basketball
- The basketball showed an orange or an apple
- The tennis ball showed a kiwi on both sides
- The kiwi showed a kiwi on both sides

After having quickly explained the task to perform to the subjects, they were invited to enter their hand in the box and explore the object. After they gave their final answers of their best guess of the objects, they were asked few additional questions.

2.4 Pilot Study

Before running the experiment, some prototypes were realised as a proof of concept. It permitted the team to adjust and define more clearly the objectives and narrow them down because of the wide range of possibilities that could be used to test the chosen modalities.
The following parameters were implemented throughout the course of the project.
Visual parameters :

- Displaying a background which was always present whether the tag was covered or not recognized (avoided the user from seeing the real object accidentally with the camera)
- Displaying a “box shape” around the object
- Having the object at a coherent size

Materials parameters :

- A bigger box
- Only having one hole (instead of two) for the hand
- Regular confirmations of the lighting setup in the box to ensure that the tag was recognized.
- Changing the tags to make them less bothering and large, so the user can feel more the object.

2.5 Variables and Factors

The independent variables (factors) were the objects put in the box and which side was presented first.
The dependant variables (measures) were the level of accuracy of the subject’s guesses, the exhibited level of confusion and the level of fun.

2.6 Data Analysis

Since the experiment was conducted with a small number of subjects, it would not be meaningful to analyse the results statistically, so the group based the research mainly on the small interviews that were conducted after the experiment to get important feedback.

3. RESULTS

In this section, the criteria, scheme and survey that were designed to evaluate the group's application were elaborated upon systematically. This survey was then applied to a set of test subjects in order to collect the necessary results and extract conclusions from them.
As previously stated in the pre-study, the focus of the evaluation of this project was not on it's psychological aspects but on the technical challenges that it entails.
Therefore, the group defined two fundamental evaluation criteria that would determine the success of the application, based on technical requirements that were to be fulfilled. These criteria were:

- Virtual consistency: (**Crit. 1**, 2nd column)
  To ensure an immersive user experience, the group set a requirement that at no point would the user perceive any object that did not belong to the virtual environment designed.
- Motion correlation: (**Crit. 2**, 3rd column)
  To ensure a smooth interaction and maximize the engagement of the user with the application, the manipulations performed by the user onto the real object were correlated with the motion of the corresponding 3D target model, with the minimum delay.

Since the objective of the game "Do you trust your senses?" was to fool the player into recognizing an incorrect object, the subjects were asked to guess the object they were manipulating after a period of 1-2 minutes of familiarization with the interface. This operation was performed twice, with different pairs of objects each time. The result of these tests are represented on the 4th (**Test 1**) and 5th (**Test 2**) columns of the table below.
Additionally, two more evaluation measures were defined, to let the test subjects give qualitative feedback of the user experience. For this, a Likert scale was used, which is a common rating scale used in research surveys [12]. Specifically, test subjects were asked to rate their user experience in terms of *Confusion*(**C**, 6th column) and *Fun*(**F**, 7th column) on a scale from 0 to 10, being 10 their maximum perceived value.
The table below shows some representative results obtained after testing the application with 3 subjects. A discussion of the conclusions extracted from these results can be found in the next section.

Subject	Crit. 1	Crit. 2	Test 1	Test 2	C	F
1	✓	✓	✗	✗	6	7
2	✓	✓	✓	✓	5	7
3	✓	✓	✗	✓	6	6

4. DISCUSSION

Even though time constraints forced a small sample size and population variance, the evaluation results are representative enough to extract a few interesting conclusions.

First of all, all evaluation criteria defined were satisfied. This means that both virtual consistency and motion correlation (explained in the previous section) were achieved. Users engaged in an immersive interaction through the application interface, and perceived the correlation between their manipulation of the real object and the simulated motion of the virtual model.
Moreover, as the evaluation table shows, test subjects guessed the correct objects with an accuracy of 50%, which indicates a significant level of disturbance in the task of object recognition.
The qualitative results are consistent with our expectations as well, since the reported ratings for the level of confusion show that all test subjects perceived difficulties while identifying the real object due to contradictory sensory information. Finally, due to the interface and user experience designs, and specially after realizing the real objects that they had been manipulating, test subjects reported their interaction as a positive and fun experience.

As a limitation of the application, the group observed that the manipulation of the real object can prevent the object detector to function properly, making it lose track of the image targets sticked to the real objects. This issue could be addressed by modifying the object detection scheme, which currently relies on a simple two-face image target. For example, obtaining a 3D model scanned from the real object would break the dependency of the object detector on an image target, allowing for a more flexible manipulation of the object, even when hand occlusion is preventing full-body detection.

One weakness detected during the evaluation process was that most users tended to rely more on the haptic modality than on the visual output from the application. In the most extreme case, one user guessed the correct object just by focusing on the touch sensory information, ignoring the visual modality. This undesired result could be corrected simply by increasing the robustness of the object detection module and replacing the current 3D virtual target models with more realistic ones, as a means to increase the immersiveness of the user experience.

Although the results show a significant level of confusion in the task of object recognition, they might be biased, as the test subjects reported that their decision could have been influenced by the assumption that the objective of the game was to fool the player. As part of the future work, this bias should be eliminated as much as possible, by, for example, alternating contradictory and non-contradictory sensory information.

5. REFERENCES

[1] Ramachandran, Vilayanur S., and William Hirstein. "The perception of phantom limbs. The DO Hebb lecture." *Brain: a journal of neurology* 121.9 (1998): 1603-1630.

[2] Petkova, Valeria I., and H. Henrik Ehrsson. "If I were you: perceptual illusion of body swapping." *PloS one* 3.12 (2008): e3832.

[3] Botvinick, M. and Cohen, J. “Rubber hands ‘feel’ touch that eyes see” Nature Volume 415, Issue 6870, 24 January 2002, Pages 429-433

[4] Ernst, M.O. and Banks, M.S. “Humans integrate visual and haptic information in a statistically optimal fashion” Nature, Volume 415, Issue 6870, 24 January 2002, Pages 429-433

[5] Srinivasan, M.A., Beauregard, G.L., Brock, D.L., “Impact of visual information on the haptic perception of stiffness in virtual environments”, Conference paper 1996

[6] Fairhurst, M.T., Travers, E., Hayward, V., Deroy, O. “Confidence is higher in touch than in vision in cases of perceptual ambiguity”, ISSN: 20452322, Published: 23 October 2018, Article number: 15604

[7] L. Hoyet, F. Argelaguet, C. Nicole, A. Lécuyer “‘Wow! I have six fingers!’: Would you accept structural changes of your hand in VR ?”, Frontiers in Robotics and AI published 12 May 2016.

[8] Pittera, Dario, Marianna Obrist, and Ali Israr. "Hand-to-hand: an intermanual illusion of movement." Proceedings of the 19th ACM International Conference on Multimodal Interaction. ACM, 2017.

[9] Lederman, Susan J., and Roberta L. Klatzky. "Haptic perception: A tutorial." Attention, Perception, & Psychophysics 71.7 (2009): 1439-1459.

[10] Loomis, Jack M., and Susan J. Lederman. "Tactual perception." Handbook of perception and human performances 2 (1986): 2.

[11] Garcia-Palacios, Azucena, et al. "Virtual reality in the treatment of spider phobia: a controlled study." Behaviour research and therapy 40.9 (2002): 983-993.

[12] Colavita, Francis B. "Human sensory dominance." Perception & Psychophysics 16.2 (1974): 409-412.

[13] Hecht, David, and Miriam Reiner. "Sensory dominance in combinations of audio, visual and haptic stimuli." Experimental brain research 193.2 (2009): 307-314.

sARcofag project
https://vimeo.com/213253969

[12] *Likert Scales and Data Analyses* by I. Elaine Allen and Christopher A. Seaman. http://asq.org/quality-progress/2007/07/statistics/likert-scales-and-data-analyses.html (Retrieved January 11, 2019).

6. APPENDIX

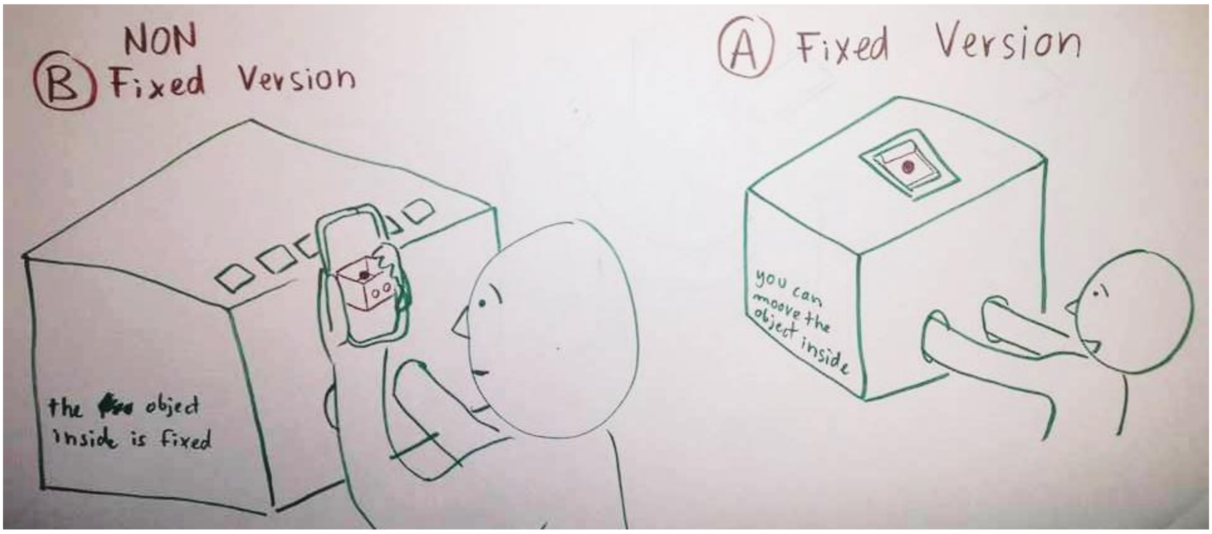

Fig 1 : part of the “storyboard” involving the two alternative versions


Fig 2 : First prototype


Fig 3 : Example of a physical object


Fig 4 : Example of a virtual object (inside the virtual box)