Virtual Subjectiveness Analysis Of The ENORASI Cane Simulator

Aarushi Sharma
Universitat Pompeu Fabra
Barcelona 08018, ES
aarushi.sharma01@estudiant.upf.edu

Albert Miquel Serrat
Universitat Pompeu Fabra
Barcelona 08018, ES
albert.miquel01@estudiant.upf.edu

Mario Tsavalos
Universitat Pompeu Fabra
Barcelona 08018, ES
mario.tsavalos01@estudiant.upf.edu

Loza Tadesse
Universitat Pompeu Fabra
Barcelona 08018, ES
loza.tadesse01@estudiant.upf.edu

Tue Ngo
Universitat Pompeu Fabra
Barcelona 08018, ES
tue.ngo01@estudiant.upf.edu

Abstract
This paper aims to give an analysis of the ENORASI cane simulator application by the means of Virtual Subjectiveness (Parés & Parés, 2006) [2]. ENORASI is a virtual environment that can be evolved into a highly interactive and extensible Virtual Reality training system for the visually impaired. We explore the context of the development of this application, the virtual experience's physical interfaces and their mappings in the virtual world, as well as the logical interface, the behaviors, and the level of interactions. We conclude that Virtual Subjectiveness is an adequate evaluation paradigm for the ENORASI cane simulator.

Author Keywords
Virtual Subjectiveness; Virtual Environment; Virtual Reality; Haptic Feedback; Auditory Feedback; Visual Impairment; Cane Simulation.

Introduction
Current Virtual Reality (VR) technologies focus heavily on visuospatial rendering and thus are inaccessible for people with limited sight or total blindness. The main objective of the ENORASI project is to tackle this problem and develop a complete training system for the blind and visually impaired that address realistic virtual representation without any visual cue from the virtual

environment (VE) (Tzovaras et al., 2002) [3]. The application was developed in the Informatics and Telematics Institute as an EU IST project. Several feasibility study tests and training scenarios have been designed, including object manipulation and cane simulation used for performing realistic navigation tasks (Tzovaras et al., 2004) [4]. In this paper, the application of cane simulation will be analyzed since it might have a direct impact on the day-to-day activities of visually challenged people in the context of commuting without potential dangers.

The authors claim that through VR, ENORASI can be used to train the visually impaired, which in the real-world would be difficult to accomplish. The training system combines haptic and aural information in such a way that helps increase the possibility of a visually disabled person obtaining an overview [3]. It assists deaf-blind and visually impaired individuals in activities such as mobility systems, accessibility of e-learning, web-based information system, and wearable interface for navigation. Specific peripheral devices and software were selected to provide a pilot environment offering adequate functionality for end-users to familiarize themselves with the technology, hence, aiding in judging the potential and usefulness of the system. The approach chosen for the ENORASI fully describes the belief of facilitating and improving blind training practices in VR and offering blind people access to new employment opportunities.

High level description

ENORASI is a VR application through which blind people can sense and understand where and how 3D objects are in a Virtual Environment (VE), thanks to two types of stimuli: force-feedback and auditory – there is no visual stimulus. In this case, there is an abstract type of Virtual Subjectiveness (VS) with a more complex structure. The VS is that of an integrated unit consisting of a virtual walking cane - an extension of the user's index finger - that can move within a virtual urban environment where there are virtual objects such as streets, pavements, and traffic lights; a virtual hand and two virtual ears positioned in the VE in such a way that they adequately perceive haptic and spatialized sound sources from the VE. It must be noted that there is no virtual person involved in the VE and that the experience is individual, not collaborative or collective.

As for the potential motion and position of the united block, the VS is constrained to navigating within the VE, as well as colliding and penetrating virtual objects. The user is able to control the VS with the movement of his/her hand and its index finger. Needless to say, the VS cannot experience any relative position of the hand to the body, as the former is the only construct present in the VE. As already established, only one hand is interchanging information with the system and that is to represent the movement of the virtual walking cane.

It is also through the use of the virtual cane that there is a fair level of VS about the presence of objects in the VE, either by sensing collision, penetration, or sensing the lack of objects thanks to the user's interpretation of the free movement of the cane. The force feedback perceived by the VS depends on the orientation of the cane relative to the virtual object that it collides with. Various force effects are interpreted as different types of collision [4]. Additionally, the VS can perceive 3D environmental sound feedback from the VE facilitating information about the relative distance of digital objects (e.g. sound beacon from a traffic light), and also

Figure 1. CyberGrasp haptic system attached to the hand of a user

Figure 2. A user performing the cane simulation.

information about collisions with them. In short, the VS can experience objects in the VE only through the haptic and aural feedback upon collisions between the virtual cane and virtual objects or by hearing sounds in the surrounding area. Amongst these objects, there are two relevant types: 1) what we could call a standing base, that is the floor on its different geometrical forms, and 2) objects that stand on the base.

The experience's goal is for the user to move freely in either an outdoor urban environment or an indoor setting and to learn further on how to use their cane and hearing to navigate properly without the dangers of the real world; or to learn about a specific mapping and configuration of the streets. Society might make some efforts to help people with special needs as the visually impaired, but an urban setting still has a long path ahead for preventing blind people from falling into the dangers faced in their daily life.

Physical interfaces

The ENORASI system consists of a variety of physical interfaces, both for the visually impaired and for the supervisor. In this analysis, we only focus on investigating those that pertain only to the VS.

The most important physical component of the system is the CyberGrasp wearable haptic device (Fig. 1) which handles both inputs from users - the motion of the hand - and output feedback to users - specific force feedback by combining a wireless CyberGlove, an exoskeleton module, and a Force Control Unit (FCU). The CyberGlove included in the CyberGrasp haptic device is fully instrumented with up to 22 sensors: three flexion sensors per finger, four abduction sensors, a palm-arch sensor, and sensors to measure wrist flexion and abduction. These sensors provide high-accuracy joint-angle measurements, and thus, the haptic device is capable of capturing precisely motion data from the fingers, hand, and wrist.

On the other hand, the lightweight, force-reflecting CyberGrasp exoskeleton is attached to the back of the hand. Grasp forces are produced by a network of tendons routed to the fingertips. There are five actuators, one for each finger and each of them can apply up to 12N force feedback. The external host PCs only communicate with the FCU through Ethernet, which in turn communicates with all other peripherals. In the particular case of cane simulation, the blind user accessing the VE will wear both the glove and the exoskeleton on one of their hands and also a waistcoat carrying the FCU (Fig. 2).

Another important physical component of the ENORASI system is the Ascension's MotionStar Wireless motion tracker which was required for tracking the position and orientation of the hand of the user. The motion tracker is a six-degree-of-freedom (6-DOF) system that uses pulsed DC magnetic fields to simultaneously track the position and orientation of a flock of sensors.

In summary, the physical interfaces of the ENORASI system include two input devices which are the CyberGrasp's 22-sensor CyberGlove and the MotionStar Wireless Tracker; and two output devices which are the CyberGrasp's exoskeleton and the host PC's speaker sending 3D auditory cues to the user. Needless to say, the user is able to receive sound feedback using their own ears, without any additional audio equipment like headphones or earphones. These physical devices restrict the area of movement to a 7m diameter.

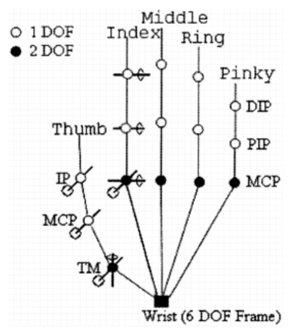

Figure 3. The DoFs of the hand model.

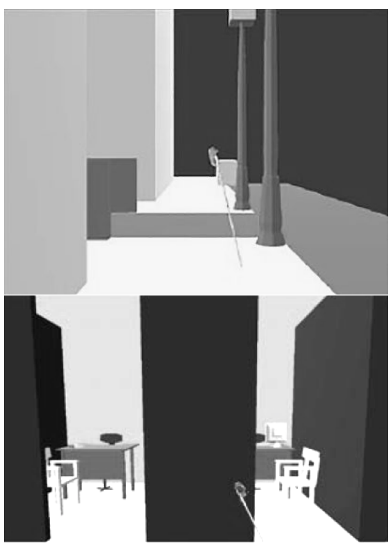

Figure 4. The virtual outdoor and indoor settings.

Logical interface

The logical interface is manifested in different global orientations (street/indoor). The combination of data obtained from the sensors of MotionStar Wireless Tracker is mapped to the position and orientation of a hand model, and the point of view is expressed in the form of this virtual hand from which the environment is constructed [4]. Each finger of the hand model consists of an inner joint, a proximal joint, and a distal joint. Fig. 3 presents the possible movements of the hand model by demonstrating the rotation axis of each joint of the thumb and index fingers. The middle, ring, and pinky fingers have the same structure as the index finger. The palm joint is located relative to the position of the tracker sensor on top of the CyberGrasp device.

After the user grasps the cane, the point of view is provided by an integrated unit including the hand model and the virtual cane itself. The virtual cane is always attached to the hand model since it is designed to be an "extension" of the users' index finger [4]. In addition, there are two virtual points resembling the user's ears which are located relative to the virtual cane in order to map the spatialized sound. Furthermore, the user is not acting through a different appearance or representation like an avatar.

Through the cane, the user is able to explore the virtual environment (Fig.4) which defines the standing direction, field of orientation, the type of projection, the mono orientation, the hearing capacities, and the force feedback properties. The parameters of the virtual cane (size, grasping, and collision forces), which handles the motion of the hand model and the specific force patterns, are adjusted so that the user feels that it is similar to the real one [4] to some extent. Although it is not explicitly mentioned in the paper, we can presume that the length and volume of the cane are likely to be predefined in order to make sense of the exact position where the tip of the cane collides with virtual objects.

The speed with which the user moves from the virtual point of view in the global orientation is adapted to the real world's speed. Walls, traffic lights, pavement, doors, chairs, or tables represent collidable objects and provide sound and haptic feedback by the system upon collision of the virtual cane with these objects.

Potential mappings

Thanks to the MotionStar wireless tracker, the real-world hand motion is translated in a 1:1 ratio to the VS' perspective in relation to the position and translation changes of the virtual hand model. CyberGrasp's 22-sensor CyberGlove collects the data related to the user's movements of the hand and maps them into the orientations and rotations of both the virtual skeleton joints and the virtual cane, which is translated in a 1:1 ratio to the VS.

The way in which the user can realize when the virtual cane is colliding with an object is through a collision detection algorithm. This is the core part of the system that ensures smooth, efficient, and accurate synchronization between the physical and logical interfaces. The researchers have evaluated the Rapid and PQP collision detection algorithms, then proposed an extension of the PQP one [4]. In order for this algorithm to work accurately and efficiently, information from both the MotionStar Wireless Tracker and the CyberGrasp's 22-sensor CyberGlove is required to determine what the virtual cane position and orientation are in the VE and to

Physical Interfaces	Data	Type	Mapping	Relation	Logical Interfaces
MotionStar wireless tracker	Position and orientation of the user's hand	Input	Orientation and direction of VS' haptic perception range	1:1	Virtual hand model and virtual cane
CyberGrasp's 22-sensor CyberGlove	22-DOF movements of the user's hand	Input	Hand's skeletal structure and joints' rotations	1:1	Virtual hand model and virtual cane
Sound speaker	3D environmental audio	Output	VS' perceived aural feedback	1:1	2 virtual hearing points (virtual ears)
CyberGrasp's exoskeleton	Haptic force feedback with different force effects	Output	VS' haptic perception: type of collision and position of collided objects	Haptic sensory substitution	Virtual cane

Table 1: Table of the different mappings

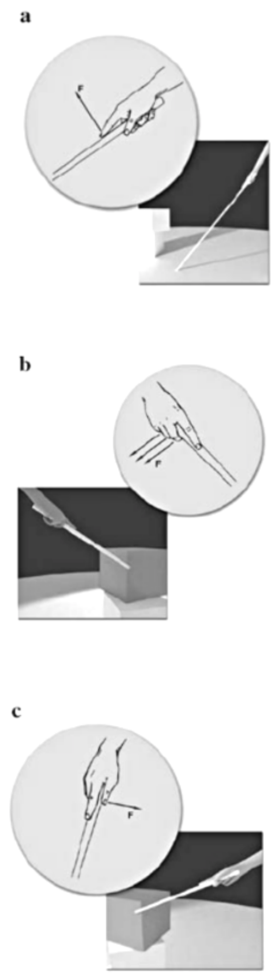

Figure 5. Cane collision with the ground (a), an object on the left side (b), and an object on the right side (c) of the virtual cane.

know where and when it enters into collision with an object. When the algorithm detects that there is a collision, two different data are mapped: auditory and force feedback. The former is in the form of 3D environmental feedback and it is provided to the VS' hearing points in a 1:1 distance ratio; the latter is also provided to the VS' haptic perception, and it varies depending on the type of the collision. The 3D environmental audio system feeds information back to the user when the virtual cane impacts with an object and also information about the relative distance of objects in the scene.

Logical interface behaviors

Thanks to the 1:1 mapping of real-to-virtual hand motion, visually impaired can explore the complex VE while freely navigating and interacting with its elements in many ways, including locating virtual objects in the VE, deciphering their sizes, and understanding their physical properties.

Since no other virtual users are present, there is no collaborative or collective activity within the VE. In fact, no other behaviors seem to exist except for the collisions between the virtual cane with virtual objects. Specifically, the force feedback applied to the hand of the user depends on the orientation of the cane relative to the virtual object that it collides with. Fig. 5 explains that when the cane hits the ground, force feedback is sent to the index finger; when the cane collides with an object laying on its right side, force feedback is sent to the thumb; and force feedback is applied to the middle ring and pinky finger in the case of a collision with an object on the left side of the cane. When the cane is penetrating an object, a cosine force effect at the buzzing level is mapped to the VS' haptic perception. On the other hand, a jolt force effect is applied to the VS' haptic perception when the cane hits the floor or a virtual object, with the maximum force being performed at the beginning and it only lasts for a short period of time.

Level of interaction

According to Parés & Parés (2001) [1], three different levels of interaction are possible: explorative, manipulative and contributive, and each includes the previous. The ENORASI project involves several applications aimed to help and train blind people, some of them are designed to manipulate objects. But in the case of the cane simulator, we believe that the type of interaction that the authors set up is explorative: the VS can explore the VE, navigate and interact with objects; however, objects cannot be rearranged or manipulated in any way. The precise aim of this application is to help the visually impaired navigate in an urban setting, not to manipulate it or to make their contributions. Hence, no rules are associated with the objects on themselves, but only with the interactions between the VS and the objects.

Interaction design approach

We believe the design and development of the ENORASI system took a content-driven approach, since the authors seemed to have a clear idea of their type of user (visually challenged people), their topic of choice (visual impaired training), and their scope (help visually impaired people recognize better their environments by practicing in a VR scenario), which let them later explore and experiment with different types of interactions. Many components of the training systems were identified beforehand, such as the necessary virtual objects for both indoor and outdoor settings (chair, pavement, traffic lights, etc.); collision detection algorithms; and specific peripheral input/output interfaces and software. Furthermore, all possible interactions between the user and the VE were scripted without allowing the user to search for or invent any metaphor, including the user's explorations and constrained manipulation of the virtual cane as well as the aural and haptic responses from the VE. Hence, we are assured that a content-driven approach was adopted.

Critiques

The cane simulator developed within the ENORASI project is a highly interesting VR application due to its lack of visual feedback. The composition of the interfaces allows the user to achieve a fair degree of VS that allows them to become immersed in the VR.

One of the author’s scopes is to train the visually impaired in indoor and outdoor navigation. The current development stage may still be prototyping, but [3] shows good feasibility results, which means the goal was achieved, at least partially. For a real-life situation training tool, the prototype is solid, but it shows several limitations related, mainly to its realism. The user can sense the virtual cane by grasping it, which facilitates a realistic feeling to the VS; however, the fact that the feedback is mapped out to the fingers as a haptic sensory substitution, even if completely coherent, lowers the degree of realism. Giving an example of CaneTroller, another haptic and aural VR application developed by Microsoft (Siu, Sinclair, Kovacs, Ofek, Holz & Cutrell, 2020) [5], there is a physical cane integrated as a part of the physical interfaces along with its representation in the logical interface. In this CaneTroller example, force feedback is mapped from the virtual cane to a wearable programmable brake mechanism that impedes the physical cane to move towards the direction where the collision occurs. This would also restrict the user from penetrating virtual objects, which, in our opinion, is a limitation of the ENORASI cane simulation experience. Furthermore, since the ENORASI virtual cane is always attached to the hand, if the user tries to hold and manipulate it in another way different from that for which it is designed, the user will quickly realize the difference between VR and physical reality.

We also think the quality and accuracy of the auditory feedback perceived (that related to the sounds emitted by virtual objects) could be improved. For example, [5] implements an HTC Vive headset for tracking head position and delivering 3D spatial audio through headphones. Additionally, it might be unnatural to get auditory feedback from a PC speaker that isn't situated where the virtual collision would occur. In that regard, the CaneTroller example shows a higher degree of realism and coherency by implementing a device placed on the physical cane that emits the simulated sound, which the VS can perceive as originated in the line between the hearing points and the virtual collision.

On a different note, the computational time required by the collision detection algorithm causes an overall delay of 10ms between the user movement and the VS perception of it. Due to this, the users have to move relatively slow in order to receive realistic haptic feedback [4]. Adding this to the fact that the physical space where users can move is limited to a 7m diameter area, the user's movements are too restricted for training in an outdoor urban setting.

Despite the above-mentioned points and taking into account the ENORASI cane simulator as a prototype, we think that the application is well-founded and its implementation is feasible: users can achieve a fair grade of VS with a very intuitive VR application.

References

1. N. Parés & R. Parés (2001). *Interaction-Driven Virtual Reality Application Design (A Particular Case: El Ball del Fanalet or Lightpools).* Presence: Teleoperators and Virtual Environments, 10, 236-245. https://doi.org/10.1162/105474601750216830
2. N. Parés & R. Parés (2006). *Towards a Model for a Virtual Reality Experience: the Virtual Subjectiveness.* Presence: Teleoperators and Virtual Environments, 15(5), 524–538. https://doi.org/10.1162/pres.15.5.524
3. D. Tzovaras, G. Nikolakis, G. Fergadis, S. Malassiotis & M. Stavrakis (2002). *Virtual Environments for the Training of Visually Impaired.* Universal Access and Assistive Technology, 1151-160. https://link.springer.com/chapter/10.1007%2F978-1-4471-3719-1_15
4. D. Tzovaras, G. Nikolakis, G. Fergadis, S. Malassiotis & M. Stavrakis (2004). *Design and implementation of haptic virtual environments for the training of the visually impaired.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 12(2), 266-78. https://ieeexplore.ieee.org/document/1304867
5. A.F. Siu, M. Sinclair, R. Kovacs, E. Ofek, C. Holz, & E. Cutrell (2020). *Virtual Reality Without Vision: A Haptic and Auditory White Cane to Navigate Complex Virtual Worlds.* 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376353