We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



2

Hapto-Acoustic Interaction Metaphors in 3D Virtual Environments for Non-Visual Settings

Fabio De Felice[1], Floriana Renna[2],
Giovanni Attolico[2] and Arcangelo Distante[2]
[1]*Computer Science department, Università degli studi di Bari "A. Moro"*
[2]*Institute of Intelligent Systems for Automation, National Research Council*
Italy

1. Introduction

Three-dimensional Virtual Environments (VEs) enable the acquisition of knowledge on a given domain through the interaction with virtual entities. The flexibility of a VE allows to represent only the part of the world that is considered relevant for the final user: the proper choice of the information, representation and rendering included in the virtual world can strongly simplify the perception and interpretation efforts required to the users. Moreover, VE can provide data that would be difficult or impossible to appreciate in the real world in an easily and simply perceivable way: domain experts can communicate specific views and interpretations of the reality in a way accessible to final users. A properly designed virtual experience can significantly improve and simplify several learning tasks.

Organizing information in three dimensions and designing techniques to interact with them require a complex effort: interaction metaphors have been introduced to facilitate the access and interaction with VEs (Bowmann, 2001). A metaphor is the process of mapping a set of correspondences from a source domain to a target domain (Lakoff & Johnson, 1980). Metaphors help designers to map features of the interaction techniques to concepts more immediately accessible to final users. Interaction can be made more immersive and engaging by multi-modality. A multimodal system coordinates the processing of multiple natural input modalities—such as speech, touch, hand gestures, eye gaze and head and body movements—with multimedia system output (Oviat, 1999). The interaction is carried out with advanced input/output devices involving different sensorial channels (sight, hear, touch, etc.) in an integrated way. Spatial input devices (such as trackers, 3D pointing devices, gesture and vocal devices) and multisensory output technologies (head mounted displays, spatial audio and haptic devices) are increasingly being used as common components of Virtual Reality applications. Each device addresses a particular sense and exhibits a different interface: (Bowmann et al., 2004) offers a broad review of multimodal interaction while (Salisbury, 2004) is a good introduction to haptics. A multimodal interaction requires data to be redundant and polymorphous to address different sensorial modalities at the same time (Jacobson, 2002).

Metaphors effectiveness strongly depends on the sensory channels they refer to and on the users characteristics. Therefore this presentation will correlate and compare hapto-acoustic

metaphors in visual and non-visual settings, evaluating their relevance when used alone or to support and complement vision. The multimodal interaction allows VE applications to address users with sensorial impairments that can rely on the particular interaction modalities they feel comfortable with. Many researches have proved hapto-acoustic VR applications to represent valid tools for visually impaired users in organizing hapto-acoustic perceptions into a mental schema of reality.

Sight is a parallel sense that allows a top-down approach to perception and cognition: it makes possible to catch a global and coarse idea of the scene in a very fast and easy way leaving to further explorations the detection, interpretation and integration of details in the mental schema of the scene. Blind mainly use touch and hear to perceive the surrounding world. Touch is a serial sense and generates a long spatio-temporal sequence of data that the user mind must integrate to build a globally meaningful mental model through a bottom-up process. Moreover, touch is ineffective or even useless when dealing with objects that are very large or that cannot be touched (far in space, sensible to damages, etc.). Contextualization can be improved by the use of acoustic cues.

Sight allows a multi-resolution approach: by using proper lenses it is easy to perceive details at scales beyond the capacity of the human eye. Touch is tied to a single level of detail (depending on the dimension of the fingertips) and cannot appreciate details that are smaller than a specific threshold. Building several proper physical artifacts (offering different enlarged versions of the objects) is expensive and can be less productive. VEs offer an easier and more effective manipulation and understanding of many kinds of information. Digital models allow the dynamic change of scale and resolution of the context to investigate with a positive impact on the human haptic perception of objects (Klatzky & Lederman, 1995) (Okamura & Cutkosky, 1999). Details of VE can be dynamically highlighted or hidden, to focus visually impaired people attention on the most relevant information.

The design of multimodal rendering must account for current technological limitations: haptic technology does not allow a stable interaction using both hands, the acoustic simulation cannot completely reproduce the richness of sounds occurring during an experience in the real world, smell and taste simulation are still far from being effective. To be preferred to tactile maps or scaled models or to the direct exploration of environment the virtual experience must be complementary with respect to the experience of reality. It must be properly created to make clear the perception and the comprehension of the desired informative content and tailored to the interests and capabilities of the user to enhance his knowledge and comprehension of the world. This poses criticism to haptic and acoustic because the combined display has to make up the lack of visual information. The limited bandwidth of haptic must be compensated by well defined interaction metaphors. This alternative way of perceiving VEs must supply a quick acquisition of the overall meaning of a scene that vision provides at a glance. Computer haptics in VEs for visual impaired is quite a recent research activity but despite to this, many different guidelines have been proposed to make hapto-acoustic interaction more effective. However there are no works reporting a taxonomy of interaction metaphors based on user tasks, as reported in the area of classic VE interaction for sighted users in (Bowmann, 2004) and (De Boeck et al., 2005).

This chapter gives an overview of the current interaction techniques in the field of multimodal VEs for non-visual interaction. The term non-visual will be used throughout the chapter to refer, further than blind users interaction, the possibility that these techniques could be used even by sighted users to support and integrate vision, when visual feedback

is not available, or unpractical, or when they cooperate with blind subjects. The proposed overview will correlate the designed non-visual interaction techniques with the well known visual interaction metaphors according to classical users tasks in VEs (navigation, selection and manipulation of virtual entities). This will make more straightforward for the reader to understand each technique and how tasks normally carried out by vision can be completed with sensorial substitution. Some of the metaphors identified as effective, integrated with new features, have been used as a conceptual framework to develop a general-purpose multimodal interaction software framework: this framework has been applied to the development of different applications experimentally verified with several blind users.
The chapter is organized as follows. In the next paragraph some reviews are analysed to better situate the taxonomy used by this presentation; then a survey of non-visual interaction metaphors will be presented and compared to visual interaction metaphors from literature. For each user task the differences between visual and non-visual setting will be outlined and the non-visual metaphors will be motivated. In section four our software interaction metaphors will be presented and in the fifth section several experiences with blind users will be reported with their subjective feedbacks.

2. Related works

In literature there are several works that propose different taxonomies of multimodal interaction techniques. In (Panëels & Roberts, 2010), (Roberts & Panëels, 2007) an overview on Haptic Data Visualization (HDV) based on data representation is presented. HDV represents the virtual model as an abstraction, encoding numerical values or an abstract mathematical concept rather a physical environment. Haptic and audio cues are used to convey analytic information in a non-visual way. This can be useful for blind users but even for sighted users in contexts where the visual feedback is not available. The taxonomy reports the following topics: charts, maps, signs, networks, diagrams, tables and images. This review outlines how most research has been performed to render charts. Other areas intensively investigated are maps, diagrams and signs.
(Nesbitt, 2005) defines a framework to support the design of multimodal displays for HDV and presents the MS Taxonomy divided in spatial, direct and temporal metaphors. Spatial metaphors describe concepts involving our perception of space and are related to spatial features such as position, size and structure. Direct metaphors describe concepts that explain the way our individual senses detect information. Temporal metaphors refer to the way we perceive occurring events in time and influence the other two metaphor types. Despite the paper proposes an articulated taxonomy, it does not go into details about possible interaction techniques.
A broad review on multimodal human-computer interaction can be found in (Jaimes & Sebe, 2007) where an overview of the field is given from a computer vision perspective. The paper focuses on the possibility of creating intelligent systems able to recognize the emotional state of the user analyzing different input sources such as facial expression and vocal recognition. In spite of the fact that haptic is just mentioned but not related systems are described, this work could help in understanding the role of haptics in affective human-computer interaction.
(Coomans & Timmermans, 1997) gives a taxonomy of human-computer communication in VR systems, focused on categorizing multimodal communication channels occurring during user virtual experiences, types of interaction and ways to display retrieved data. No classification of adoptable techniques according to user task is given.

This chapter addresses interactions metaphors according the user tasks and the metaphors proposed to effectively accomplish them. User tasks, in our opinion, are a more general and straightforward way to ground interaction techniques; they can be considered as events occurring in time therefore they can be linked with temporal metaphors (Nesbitt, 2005). In particular navigation metaphors can be applied to maps as described in the works of Panëels and Roberts. We can refer to selection and manipulation as general interaction tasks, with which the user picks up an object and manipulates it both in a physical sense (translating, rotating or scaling it) and in a semantic sense (extracting enclosed information). The latter manipulation technique is particularly related to techniques for data visualization therefore manipulation metaphors can fall inside spatial and direct metaphors and can be connected to techniques used for charts and maps.

3. Hapto-acoustic interaction with Virtual Environments

The classification of users tasks in VEs we refer to comes from the works of Gabbard, (Gabbard, 1997) and Esposito (Esposito, 1996) that organize them in:

- Navigation
- Object Selection
- Object Manipulation, Modification and Querying

Based on this classification many interaction metaphors have been designed and implemented. The majority of them are based on vision as the main feedback channel.
In this section, a survey of the most used metaphors for each task will be given, their hapto-acoustic version, if possible, will be described with their relevance in non-visual settings.

3.1 Navigation metaphors

The navigation task refers to how users move inside the VE. Many works (Bowmann, 2005), (Shermann & Craig, 2003) agree in distinguishing two main components of the navigation task: travel and wayfinding. The former is the navigation physical component: the user moves through a series of locations exploring the environment, investigating its structure and the position of virtual entities. Other travel sub tasks can be the search tasks (moving to a specific target location) or manoeuvring tasks (highly precise small movements place the user viewpoint in the position/attitude most suitable to a specific goal). Wayfinding is the cognitive component of navigation: it involves the synthesis of the information acquired while travelling into a mental representation of the space and its use to plan paths in the VE.

3.1.1 Physical movement based metaphors

A group of travel metaphors is based on user physical movements. These metaphors use the body motion to drive the user movement inside the VE: most of them are intended for immersive VEs (such as CAVE applications). Types of physical metaphors are:

- Walking
- Walking in Place
- Devices Simulating Walking
- Cycles

Walking is the most direct form of travelling in a 3D world. It is a natural technique and provides vestibular cues but technological and spatial limitations make it not always feasible. It can be used when the VE size is smaller than the tracker working area. Other

solutions (Welch et al., 1994) use an optical tracking system that enables the use of a wider area by using a scalable tracking grid on the ceiling. Walking is more important for 3D applications such as mobile Augmented Reality (Höllerer et al., 1999). Walking in place is a natural alternative to walking: users move their feet to simulate walking always remaining in the same place. It does not have spatial limitations and still allows users to drive the movement with their body but it does not provide vestibular cues and gives a lower sense of presence in the VE. Devices Simulating Walking metaphors can be applied in very large VEs: they simulate the real walking with various types of treadmills allowing a more effective walking but they are expensive and do not provide the perception of a natural walking. Rather, user learns how to adapt his walking motion to the device characteristics. Finally, Cycles devices can be applied to simulate walking: they are less effective in providing a real walking sensation but are less complex and expensive than treadmills.
Body based haptic devices, such as vibrotactile gloves, belts or suits (Piateski & Jones, 2005), can be integrated in the walking metaphor and can define useful tactile cues to convey information even for blind users. In particular, vibrotactile solutions for non-visual autonomous mobility are currently investigated inside the (HaptiMap European project).
Ground based haptic devices, such as the PHANToM or the CyberForce, could be suitable for integration with other physical metaphors both in contexts of visual and non-visual feedback. The whole hand haptic interaction can be useful to feel the grasping of virtual objects while stylus based haptic devices can simulate the blind cane. However, to our knowledge, there is no literature in this direction.

3.1.2 Steering based metaphors

Steering is the continuous specification of absolute or relative direction of travel in different ways and with different devices. These metaphors share a direct camera manipulation approach. The techniques we will describe are:

- Gaze-directed
- Pointing
- Flying Vehicle
- Scene in Hand
- Eyeball in Hand

When a head tracker or an eye tracker is available the most natural choice for travelling in the VE is the Gaze-directed steering (Mine, 1995). With this metaphor, the direction the user is looking at specifies the direction of motion. In spite of its simplicity, this solution can become inefficient because whenever the user looks around to explore the environment he changes the motion direction. To avoid this problem the Pointing technique has been introduced in (Mine, 1995). It is based on two separate trackers: one used to look around in the VE while the other one, generally held in the user hand, specifies the direction of travel.
(Lécuyer et al., 2003) presents a system for blind people based on the egocentric navigation: travelling is constrained by a predefined path where the user moves using a wireless gamepad while a haptic device simulates a cane used to feel the VE. The exploration is integrated with auditory and thermal feedbacks (providing a perception of the sun position) besides the visual one, that allows sighted people to follow the navigation. This system can be Considered as using the Pointing Metaphor, where the trackers are respectively the gamepad, used to navigate the VE, and the cane used to look around. Criticisms mainly concern the constraint posed on the route and the constant speed of navigation; moreover the thermal effect needed additional work and the virtual cane was not always perceived.

In (Ware & Osborne, 1990) three metaphors are described and compared. The first is the Flying Vehicle metaphor where the virtual camera is represented as mounted on a virtual vehicle and the user can move the vehicle specifying its direction of motion. This is the most used travel technique applied in immersive and desktop solutions whose major drawback is the time required to travel between two far apart points in large VEs, depending on the maximum allowed speed. In the Scene in Hand metaphor, a virtual object is taken as an anchor around which the user can orbit: the direction of movement is related to the object position. This is useful for manoeuvring tasks while it results less effective for exploration and search tasks. The third metaphor is the Eyeball in Hand: the user holds a 6DOF tracker in his hand as if he would really hold his eyeball in the hand whose movements are directed coupled with the camera motion in the virtual space.
The Flying Vehicle metaphor can be extended to haptic display in the context of visual feedback (Anderson, 1998). Haptic feedback allows feeling objects through vibration and bump, according to the type of haptic device at hand.
In (De Boeck et al., 2001) and (De Boeck et al., 2002) the Camera in Hand metaphor is presented as a hapto-acoustic extension to the Eyeball in Hand technique. This technique uses the PHANToM haptic device to allow the user to combine Pointing and travel tasks in a single device avoiding the use of any further distracting tools. The virtual camera is directly coupled with the stylus of the haptic device and the user can specify camera direction and orientation in an absolute manner by moving the haptic stylus. Further enhancement to this technique (Extended Camera in Hand) allows switching from the absolute mode to a relative mode of direction specification to explore VEs larger than the physical haptic workspace. Acoustic feedback is used to highlight the hardness of pushing and the subsequent navigation velocity.
Although Flying Vehicle and Camera in Hand metaphors allow the use of haptic and acoustic cues in visual setting, they are not effective in non-visual setting. For sighted users, the point of view of the VE is the perspective from the position of the virtual camera. In hapto-acoustic applications addressing blind or eyes busy users, the point of view can be defined as the perspective from the position of the haptic probe. Directly coupling the haptic stylus with the virtual camera enables an egocentric point of view of the VE and a subjective navigation. In this type of navigation, the point of view continuously changes depending on the user hand movements and this can confuse blind users that lack the stable references points provided by vision (Sjöström, 2001). Indeed, this missing visual feedback makes more difficult to match information acquired from different point of views and very hard to construct a global schema of the VE.
Steering metaphors with direct camera manipulation, involving a continuously changing point of view, can be integrated with hapto-acoustic cues to enhance the sense of immersion of sighted users. In case of visually impaired or eye busy users, the use of Steering metaphors with haptics as the main feedback channel can be confusing.

3.1.3 Route planning and Target based metaphors

Another way of travelling is through the specification of a travel route or of the given target. Both these approaches have in common an indirect camera manipulation because the viewpoint is automatically manipulated by the system after the specification of the route or of the final destination made by the user. Route Planning metaphors constraint the navigation along the path specified by the user: the system handles the sequence of

movements. The user has a minimum control on the travel and can pay attention to other tasks. Described metaphors are:

- Drawing a Path
- Marking Points

The Drawing a Path metaphor allows the user to draw a 2D path that is then projected into the 3D space. In (Igarashi et al., 1998) the 2D line is specified using the mouse. The Marking Points metaphor requires the placement of markers along a path of the VE: the system is responsible for creating a continuous path through this sequence of points. In (Bowmann et al., 1999) the user places markers using a 3D map of the VE; points marked on the map are then projected on the VE and connected by straight segments. An advantage of this technique is the possibility of changing the precision of the path by placing a greater or lower number of markers.

If the entire route is not relevant, it is possible to directly move the viewpoint to a target specified by the user. Some target based metaphors are:

- Teleportation
- Small Scene Manipulation
- Map-based Specification

With Teleportation users are instantly brought into a given position in the scene. This approach can easily disorient sighted users (Bowmann et al., 1997) because jumping from one place to another significantly decreases spatial orientation. A gradual movement from the starting point to the target is always recommended to reduce this effect. The Small Scene Manipulation metaphor has been introduced in (De Boeck et al., 2004): the camera smoothly zooms in a particular part of the world chosen by the user. Map based specification employs a 2D map depicting the entire world over the 3D VE and users can trace markers everywhere in the scene without being constrained by the viewpoint. The World in Miniature (WIM) metaphor (Mine, 1996) implements this approach allowing the user to place markers by manipulating a 3D miniature version of the VE. This metaphor can be considered more than a navigation metaphor because users can even select and manipulate distant objects through their miniature representation.

Haptic integration in Route Planning and Target based techniques is not well documented for visual settings. Moreover it is judged not useful for indirect camera manipulation in visual feedback context (De Boeck et al., 2005). Something analogue to a haptic route planning approach can be found in (Vidholm et al., 2004). This paper investigates how stereo graphics and haptics can be combined to facilitate the seeding procedure in semi-automatic segmentation of magnetic resonance angiography images. A force feedback based on local gradients and intensity values generates constraint forces allowing users to trace vessels in the image by placing in the 3D data set the seed points that will guide a region based segmentation algorithm.

On the other hand, planning routes and defining targets in VE are very useful techniques that designers can employ to develop hapto-acoustic metaphors for virtual navigation of visual impaired. For this kind of users, the combination of attractive forces and magnetic paths can implement a sort of haptically guided tour through the VE (Sjöström, 2001), very useful to support non-visual wayfinding.

Different solutions are available for the haptic guided exploration. (De Felice et al., 2007) and (Pokluda & Sochor, 2005) propose an exploration based on free movements with guide: the user can freely explore the VE, asking for being guided to a known point whenever he feels lost. Other applications fully constrain user movements to follow a predefined path (Roberts et

al., 2002), (Pokluda & Sochor, 2003). In (De Felice et al., 2005), blind can switch from a free exploration of a 3D reconstruction of an ancient pillory to a guided tour through the most interesting features of the model (Figure 1). Whenever to user is brought in a target, a synthetic speech informs him with historical data of that feature. The guiding force must be gently enough not disturbing the user attention but sufficiently strong to be perceived and followed. In this work the following logarithmic force feedback has been used:

$$F_a = \mathrm{Log}(D) - \varepsilon \quad (1)$$

where D is the current distance between the haptic probe and the target, while ε is an offset taking into account the fact that the 3D position of the target is placed inside the object mesh. The force direction is given by the vector connecting the haptic probe to the target. The force decreases as the haptic probe comes closer to the used feature: the user can freely orbit around the target being always haptically advised to come closer to it whenever he moves too far. In non-visual navigation, routes and markers are needed to supply the haptic experience with tools that enhance the construction of the global schema of the VE. These types of planning must be predefined by designers in the non-visual contest while they can be dynamically specified during the navigation by sighted users.

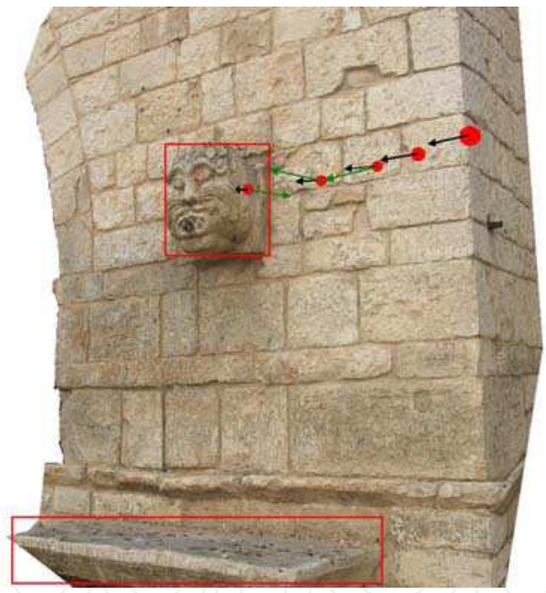


Fig. 1. A haptic guided tour on a 3D model. The sphere is attracted by the salient portions of the model (rectangles) while a synthetic speech gives information on historical data. Black arrows depict the attractive force, while the green arrows describe the contact force.

The World in Miniature metaphor offers an effective starting point to describe the most diffuse approach for navigating VE in application with non-visual display. As stated above, a subjective navigation involving a continuous change of the viewpoint is not suitable for blind navigation, and VE to result effective must overcome the traditional difficulties that blind subjects experience in the real life. For these reasons, many applications propose the entire VE as a world in miniature that fits the workspace of a one-point haptic device. In this way, the blind user has the entire world below his hand and can navigate rapidly and safely from one point to another (Figure 2). The movement of the stylus does not affect the

reference system of the VE, that keeps its absolute position and orientation in relation to the desk. The avatar moves in a physical workspace, depending on the haptic device, that the user explores in an exocentric way.
Many works refers to this setting when VE helps blind users to navigate in real environments (Magnusson & Rassmus-Gröhn, 2004), (Lahav & Mioduser, 2008), (De Felice et al., 2007), (Murai, 2006) or to acquire geographical data from virtual maps (De Felice et al., 2007), (Parente & Bishop, 2003), (Moustakas et al., 2007). Experiments with VE representing real environments investigate the possibility that a preliminary virtual exploration can support the construction of an effective cognitive map, helping blind to navigate the real counterpart in a more autonomous and secure way. (Magnusson & Rassmus-Gröhn, 2004) used the PHANToM to virtually navigate a complex traffic VE. The study shows that most of the users were able to understand and navigate the VE and that users with better performances in the VE were very good at moving with a cane. Moreover, it outlines that blind users require detailed and articulated 3D virtual worlds so that it is important to find more automatic ways to create such environments. (Lahav & Mioduser, 2008) developed a 3D multimodal indoor VE that can be navigated with the Microsoft Force Feedback Joystick. They compared the performance in the real environment of people trained on the VE with those of the control group that directly explored the real one. They state that VE allows users to navigate real environments better than people belonging to the control group because virtual navigation allows a comprehension of all the parts, even the inner ones, that is faster and safer. (Murai et al., 2006) used the PHANToM Omni to explore an indoor environment. The use of a virtual simulator can offer some remarkable features such as guidance by using two haptic devices, one for the trainer and one for the user. It also reported that an extended training with the application enhances users performances and the system efficacy compared to traditional media. In (Yu & Brewster, 2002) a system designed to improve visually impaired people access to graphs and tables is presented.

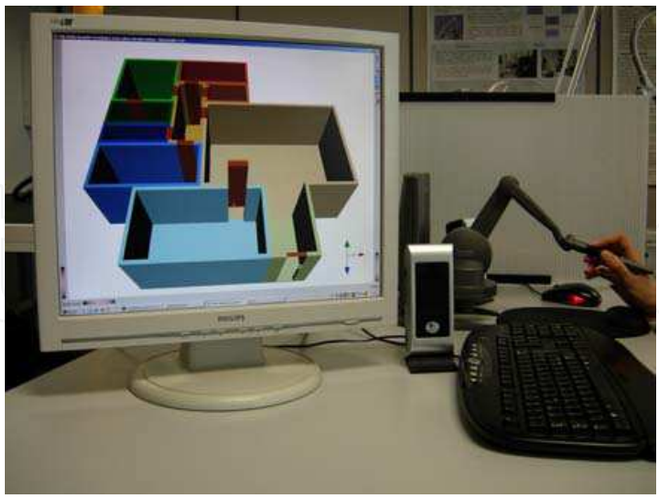


Fig. 2. The hapto-acoustic version of the World in Miniature metaphor used to navigate VE. The user pilots the avatar through the haptic stylus of a PHANToM Desktop

Table 1 reports an overview of the described metaphors highlighting their compatibility with haptics in visual and non-visual setting. Their typologies are reported as pm (Physical Movement), s (Steering), rptb (Route planning and Map Target based).

Metaphors	Metaphor type	Other tasks possible	Compatible with haptic (visual)	Compatible with haptic (non-visual)
Walking in place	pm	no	yes	possible
Treadmills	pm	no	yes	possible
Cycle	pm	no	yes	possible
Gaze Directed	s	no	no	no
Pointing	s	no	no	no
Flying Vehicle	s	no	possible	possible
Scene in Hand	s	no	possible	yes
Eyeball in hand	s	no	no	no
Camera in Hand	s	no	yes	no
Teleportation	rbtb	no	no	no
Small Scene Manipulation	rbtb	no	no	yes
World in Miniature	rbtb	selection/ manipulation	possible	yes

Table 1. Navigation metaphors overview. Compatibility is possible if haptic can be useful but little or no related literature is available.

3.2 Selection metaphors

Selection, also called target acquisition task (Zhai et al., 1994), refers to the acquisition or identification of a particular object in the virtual scene. Selection techniques, that require the indication of a particular object and the confirmation of its selection, can be grouped in:

- Virtual Hand
- Ray-casting
- Aperture-based
- Speech

Virtual Hand, the most intuitive and used selection technique, allows the user to directly select and manipulate an object with his hand. Ray-casting metaphors allow the object selection by pointing a ray at it. These mentioned metaphors have hapto-acoustic extension for visual and non-visual settings, and will be described in the following sub sections.

Aperture-based metaphors (Forsberg et al., 1996) provide users with an aperture cursor sliding over a cone shaped pointer that enables the dynamic control of the cone width whose apex is at the user eye and of the central axis whose direction is specified by the hand tracker movement. At our knowledge, no works have addressed the multimodal enhancement of these metaphors in non-visual settings.

The Speech metaphor (De Boeck et al., 2003) is a very natural and intuitive solution when selectable objects can be univocally named but in complex environments disambiguation can be required and the Automatic Voice Recognition applications are still error-prone. The integration with force feedback is difficult. Sound cues can be used to confirm selections.

3.2.1 Virtual Hand

With the Virtual Hand metaphor, the user moves a virtual cursor by using a tracker such as the Cyber Glove or any 3 DOF input device. The virtual cursor can have different shapes, typically it is a human hand 3D model. The user hand motion is directly mapped to the virtual hand motion by calculating its 3D position $\mathbf{p}_v$ and orientation $\mathbf{R}_v$ in the VE:

$$\mathbf{p}_v = \alpha \mathbf{p}_r \; ; \quad \mathbf{R}_v = \mathbf{R}_r \tag{2}$$

where α is a scaling factor while $\mathbf{p}_r$ and $\mathbf{R}_r$ are position vector and orientation matrix of the user hand respectively. Whenever the virtual hand intersects a virtual object, it becomes selected. The movements of the selected object are directly coupled with the user hand motion, making this metaphor also a useful technique for manipulation tasks (for details see section 3.3). This metaphor results very intuitive and natural because allows a behaviour similar to touch an object in every day life. Its major drawback is the limited workspace into which objects can be selected, due to the limited range of user limb, making a distant object only selectable after a navigation task. To overcome this limitation, in (Poupyrev et al., 1996) the Go-go metaphor has been proposed, which provides to the user a way to interactively and non-linearly change the length of the virtual arm.

3.2.2 Ray-casting

Ray-casting is by far the most popular distant selection metaphor that allows the user to select objects by directing a virtual ray. In immersive solutions the ray can be attached to the user virtual hand, while in desktop solutions can be attached to a 3D widget controlled by a mouse. The basic version of the metaphor allows to select the closest object intersecting the ray, resulting in a very effective and simple local selection technique while its precision significantly decreases selecting far objects (Poupyrev et al., 1998), (Bowmann et al., 1999), (De Boeck et al., 2004). Indeed, when the ray operates over large distances it results in a lower accuracy due to the hand jitter amplification; moreover small angular motion of the input device causes large movements of the ray. Objects occluded or in high dense arrangement are difficult to be selected (Liang & Green, 1994). The Cone-casting enhances the basic technique involving the use of a cone shaped selection volume with the apex placed at the virtual input device 3D position (Liang & Green, 1994), (Zhai et al., 1994). By increasing the activation area, multiple targets can be selected and a more stable pointing is allowed but this approach, due to the need of a disambiguation mechanism, leads to a more complex interaction (Hinkley et al 1994).
(Vanacken et al., 2009) presents and compares the 3D Bubble cursor, Depth ray and Haptic Lock ray techniques trying to overcome the ray-casting metaphor limitations; the visual and multimodal versions of each metaphor are tested in environments with different degrees of density. To highlight selection and confirmation of targets, 3D Bubble cursor and Depth ray multimodal versions introduce short sinusoidal bumps and earcons (Cockburn & Brewster, 2005). The multimodal feedback for Haptic Lock ray is dynamically activated when the user disambiguates multiple selected targets: the haptic probe behaves as a depth marker, haptically constrained along the ray vector and whenever it intersects an object, the user feels a haptic bump and an earcon. Experiments showed that the added multimodal feedback did not provide any real improvement in planning the selection of occluded objects in dense scenes while the advantage of hapto-acoustic cues is in increasing the user awareness of the target capturing operation. The use of haptic with stereo graphics for the

selection task is reported in (Wall et al, 2002). Users had to select targets in a 3D VE in different experimental setups switching from 2D visual feedback to stereo visual feedback, with and without haptic feedback. Results suggested that the haptic feedback provides insignificant benefits in enhancing target selection time especially when stereo graphics is used and are consistent with (Akamatsu et al., 1995), (Cockburn & Brewster, 2005) works.
In non-visual setting, the selection task is a quite passive procedure due to the lack of global knowledge of the whole scene; active non-visual selection techniques should provide both a way to show the user the objects available for the selection and tools to select them and to confirm the choice. (Magnusson & Rassmus-Gröhn, 2005) compares a set of hapto-acoustic tools to help users in selecting objects within a memory game. In particular, attractive forces and Linear Fixture (Prada & Payandeh, 2005) have been compared. Linear Fixture constrains the haptic probe on a magnetic line pointing towards the selected object where the user can move. Results showed no preferences between the two haptic techniques and that user ratings depend on the type of interaction. Moreover, 3D positional sound was preferred chosen among the proposed acoustic designs. In (Ménélas et al., 2010) hapto-acoustic metaphors are used to enhance target selection in non-visual VE with multiple and occluded objects. The Virtual Magnet metaphor is used: an attractive sphere placed around targets attracts the haptic probe with a force decreasing as long as the probe approaches the target. The positional information (via spatialisation) and distance (via repetition rate and level variations) are the two proposed acoustic cues. Comparisons are made between acoustic only, haptic only and multimodal conditions resulting in a lower selection time with haptic only and multimodal conditions.
An approach to increase the activation area allowing multiple selection candidates is presented in (Pokluda & Sochor, 2005): the selection area is a 3D sphere (Navigation Sphere) centred in the user avatar position that limits the haptic device movements to a small radius. Objects in the scene are projected on the sphere surface and the user can perceive distal virtual objects by haptic effects (vibrations). In our opinion, this is a very useful technique for selection in non-visual setting because it allows the acquisition of a global haptic view of the available objects at any distance. Unfortunately the paper lacks detailed blind users feedback. An interesting extension could be obtained integrating the Navigation Sphere technique with a guiding force moving the user hand toward the object.
In table 2 an overview of the selection metaphors is reported highlighting compatibilities with hapto-acoustic cues in visual and non-visual settings.

Metaphors	**Distant action possible**	**Other tasks possible**	**Compatible with haptic (visual)**	**Compatible with haptic (non-visual)**
Virtual Hand	no	Manipulation	yes	possible
Go-go	yes	Manipulation	yes	possible
Ray-casting	yes	Manipulation	yes	yes
Cone-casting	yes	Manipulation	yes	yes
Apertured-based	yes	no	possible	possible
Speech	yes	no	no	no

Table 2. Selection metaphors overview

3.3 Manipulation metaphors

Manipulation in VEs consists in translating, rotating and scaling the selected objects to acquire, through sight, a full knowledge of their shape and geometry for general interaction purposes. According to (Poupyrev et al., 1998), the manipulation task can be divided into exocentric and egocentric manipulation. In the former the user manipulates objects acting from outside the world with a god eye's view while, in the latter, he acts from inside the world and manipulates objects with a subjective point of view. Multimodal manipulation allows an enhanced examination of VE in visual settings because more properties can be acquired: shape and geometry with sight, material and weight with haptic, acoustic properties with hearing. In non-visual settings, touch is the main feedback channel and manipulation must supply features to make effective haptic examination of objects for a complete comprehension of their features. Egocentric manipulation metaphors includes:

- Virtual Hand
- Object in Hand
- HOMER
- Image Plane

while exocentric manipulation are:

- World in Miniature
- Scaled World Grab
- Voodoo Dolls

3.3.1 Egocentric manipulation

Egocentric manipulation metaphors allow to interact with the world from a first person viewpoint. In visual settings, these techniques are well suited for manipulation tasks such as object deformation, texture change and 3D menu interaction. In non–visual settings, egocentric manipulation can help as long as the examined object is not moved, otherwise reference points are missed confusing the user.

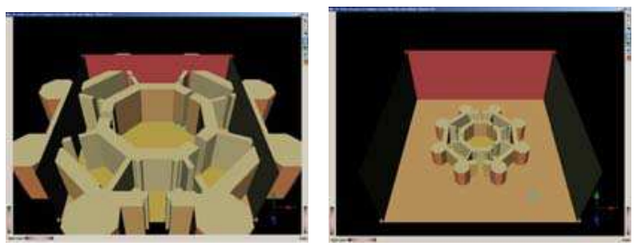
Direct manipulation metaphors such as the Virtual Hand and Object in Hand can be integrated with haptics and acoustic feedbacks. HOMER (Hand-centred Object Manipulation Extending Ray Casting) metaphor (Bowmann & Hodges, 1997) allows to select objects as in Ray-casting technique; then the user virtual hand moves to the object position where a direct manipulation can be applied. Despite the lack of works specifically addressing the integration of this technique with hapto-acoustic cues, it is plausible that multimodal versions of this metaphor could be obtained mixing the Ray-casting hapto-acoustic version and the virtual hand metaphor haptic version. Image Plane technique (Pierce et al., 1997) allows the user to select, move or manipulate objects indicating their projection on a 2D screen with a 2D mouse; since this is a 2D interaction technique for a 3D world, it is not possible to have a 6 DOF manipulation so haptic is not an added value.

3.3.1.1 Virtual Hand

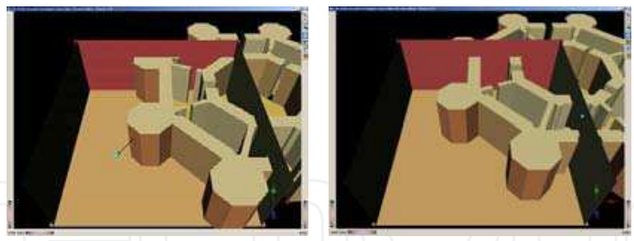
As already stated in 3.2, Virtual Hand allows users to directly select and manipulate objects in the VE. Once the object has been selected, virtual hand movements are directly applied to the object in order to move, rotate and scale it.

This metaphor is well suited for the integration with the haptic feedback. It can be used with one-point haptic devices, such as the PHANToM Desktop from the (Sensable), and whole hand devices such as the Haptic Workstation from the (Immersion). Whole hand haptic

manipulation in visual environments can be very useful in application of virtual prototyping and CAD design. (Ott et al., 2010) addresses the possibility for the user to interact with objects with both hands and the Haptic Workstation is used. The enhanced realism and effectiveness in VEs allowed by the haptic feedback require a greater complexity in computing forces for one or two whole hands manipulation with respect to a one point interaction. They propose a haptic hand model based on the God-object method of (Zilles & Salisbury, 1995), a proxy-based method largely used for one-point interaction devices. The virtual hand is decomposed in a Tracked Hand, directly coupled to the user real hand data; the Proxy that is a mass-spring-damper system connected to the Tracker hand by a set of viscoelsatic links and, the Visual Hand that is the only visible one. As long as the user moves into free space, the Proxy follows the Tracked Hand and the Visual Hand reflects the Proxy configuration. Whenever a virtual object is intersected, the Tracked Hand is allowed to penetrate its geometry, while the Proxy and consequently the Visual Hand are constrained to remain on the object surface. The viscoelastic links, connecting the Tracked Hand and the Proxy, determine the force feedback magnitude and directions to be applied to the device. Other works addressing the haptic rendering of the whole hand are (Garre & Otadui, 2009) for the manipulation of deformable objects and (Tzafestas, 2003) that also investigates the problem of torques rendering. They do not refer to any acoustic feedback.



(a)



(b)

Fig. 3 Haptic manipulation techniques: a) zooming in and out, b) PHANToM dragging and Box dragging.

Haptic manipulation in non-visual settings aims at understanding objects rather than moving them around changing their position and orientation. Applications are mainly based on one-point interaction devices, that are more stable and effective if compared to whole hand armatures that resulted still not very effective in contexts where haptic is the only feedback channel. Moreover, one-point devices allow to easily add different haptic effects such as vibration and attraction using SDKs as Chai3D (Conti et al, 2005) or H3D (H3Dapi) while whole hand devices need to program these effects from scratch. Simulating special haptic effects, further than the basic contact force, supplies different ways to acquire

knowledge about shape, geometry and details. In (Kolcárek & Sochor, 2005) a velocity driven LOD is introduced to allow blind users to directly manipulate and examine 3D virtual objects at different level of details. The algorithm is based on the assumption that fast movements of the haptic probe correspond to a user intention of acquiring a coarse idea of the object, while slow movements correspond to user intention of obtaining more details. Consequently, the 3D mesh is rendered at a simplified version for fast probe movements, while more detailed versions are dynamically rendered as speed slows down. Seven out of ten subjects participating in the test stated that the method effectively helps to better recognize global shape and details of the virtual objects. Other type of features, that can be added to object in order to better comprehend them, are dragging and scaling techniques. In (Magnusson & Rassmus-Gröhn, 2003) dragging and scaling functionalities are provided to dynamically change the part of a large model to be shown in the workspace and its scale. Three types of dragging are given. The PHANToM dragging technique gives the user the sensation of moving the whole model of the VE (according to the movements of the stylus) with respect to the position of a containment Box. In the box dragging technique the user seems to move the containment box (by pushing on its walls) with respect to the VE. Both these techniques are depicted in Figure 3b. Finally, the keyboard technique allows users moving the world pressing the arrow keys. The scaling functionality, depicted in Figure 3a, dynamically changes the sizes of scene details according to the dimension of the user avatar. All these operations are highlighted by a suitable sound (a sliding rock for dragging, increasing and decreasing sounds for zooming in and out). Users used without preferences the three dragging techniques and judged intuitive the zoom technique.

3.3.1.2 Object in Hand

The Object in Hand metaphor has been developed by (De Boeck et al, 2004) in order to exploit both user hands during the interaction with virtual objects: the non dominant hand can be used to grab and bring directly selected objects into a comfortable position to be better manipulated by the dominant hand. This metaphor is directly implemented with haptics enabled, indeed a CyberGrasp is used to grasp the virtual object. When the manipulated object is released, it comes back to the initial position. The main advantage of this technique is its intuitiveness that allows users to manipulate objects as in every-day life while its main drawback is the duplication of devices with their burden of cables. Its use in a non-visual settings is not documented, although it could be interesting to investigate how blind users can take advantage from this type of interaction taking into account that the returning of the manipulated object to its initial position after the manipulation can confuse them.

3.3.2 Exocentric manipulation

With exocentric manipulation, users are able to manipulate objects from a god-eye viewpoint, selecting and manipulating them everywhere in the VE, without being constrained to the local area as with egocentric techniques.

World in Miniature metaphor can be used for selecting and manipulating tasks further than navigation. Users can reach very quickly any object in the environment without navigation efforts. Haptics and acoustic cues provide direct and intuitive feedback on the accomplished task; this can be especially useful for blind or eye-busy users. The main drawback is the lack of accuracy due to the small scale of the miniature representation.

With the World-scaled grab metaphor (Mine & Brooks, 1997) the user can select and manipulate distant objects by pointing them with the Ray-casting, then the entire VE can be

scaled with respect to the user viewpoint in order to make the remote selected object reachable by the user hand. The manipulation is then possible as with the virtual hand metaphor but with the disadvantage of having smaller object dimensions and a less accurate manipulation. The scaling factor α_s is calculated as follows:

$$\alpha_S = \frac{D_V}{D_O} \tag{2}$$

where D_v is the distance between the user viewpoint and the position of the virtual hand and D_o is the distance between the user viewpoint and the selected object.
The Voodoo Dolls metaphor (Pierce et al, 1999) allows manipulating distant objects in an exocentric frame of reference. It is a two hands interaction technique usually implemented with hand trackers such as the Pinch Glove or the CyberGlove. The user can select an object and create a miniature version of it, called doll, with the non-dominant hand. This doll represents a stationary frame of reference, meaning that if the user moves the doll, the corresponding distant object does not move. The user can select another object with his dominant hand while its position and the orientation in the stationary frame of reference are defined by the created doll.
Consideration about hapto-acoustic integration for these techniques are analogue to those reported for the Virtual Hand metaphors because they both allow a direct manipulation. No relevant literature can be found about manipulation techniques in non-visual settings. It is plausible that much work must be done to systematically observe the behaviour of blind or eye-busy users engaged in articulated multimodal manipulation tasks based on these metaphors.

Metaphors	Distant action possible	Other tasks possible	Compatible with haptic (visual)	Compatible with haptic (non-visual)	Taxonomy
Virtual Hand	no	Manipulation	yes	possible	Egocentric
HOMER	yes	Selection	yes	possible	Egocentric
Object in Hand	yes	no	yes	possible	Egocentric
Image Plane	yes	Selection	no	no	Egocentric
World in Miniature	yes	Selection/ Navigation	yes	yes	Exocentric
HOMER	yes	Selection	yes	possible	Exocentric
Voodoo dolls	yes	no	possible	possible	Exocentric

Table 3. Manipulation metaphors overview

4. The OMERO framework

The most effective techniques to accomplish user tasks, taken from literature and experimented in our past works (De Felice et al., 2005) and (De Felice et al., 2007), have been chosen and integrated to develop a multimodal interaction framework called OMERO (Organized Multimodal Experience of Relevant virtual Objects) that supports haptic devices characterized by one-point interaction and kinesthetic feedback and users tasks as navigation, objects selection, objects manipulation and scene querying. We introduced the

Scene Querying as a task to dynamically display the scene informative contents in response to particular user queries. This task is supported by the Scenarios metaphor. Moreover, we defined Active Objects as a metaphor to support object selection and manipulation. A conceptual framework has been developed to integrate the interaction techniques, (Figure 4). Based on this conceptual model a set of multimodal applications, described in the following section, has been developed and tested with groups of blind and low vision subjects.

4.1 The Active Object metaphor

A VE can be defined as a set of virtual objects (Figure 4). A given virtual object is described through its name, geometry and haptic material. A distinction is made between Active and Background objects. The former are objects with which users can interact by means of selection and manipulation while the latter convey to the user only their shape. Active Objects can be associated with simple behaviors (haptic, acoustic or dynamic) or complex behaviors that are combinations of the simple ones. Haptic behaviors are given by haptic

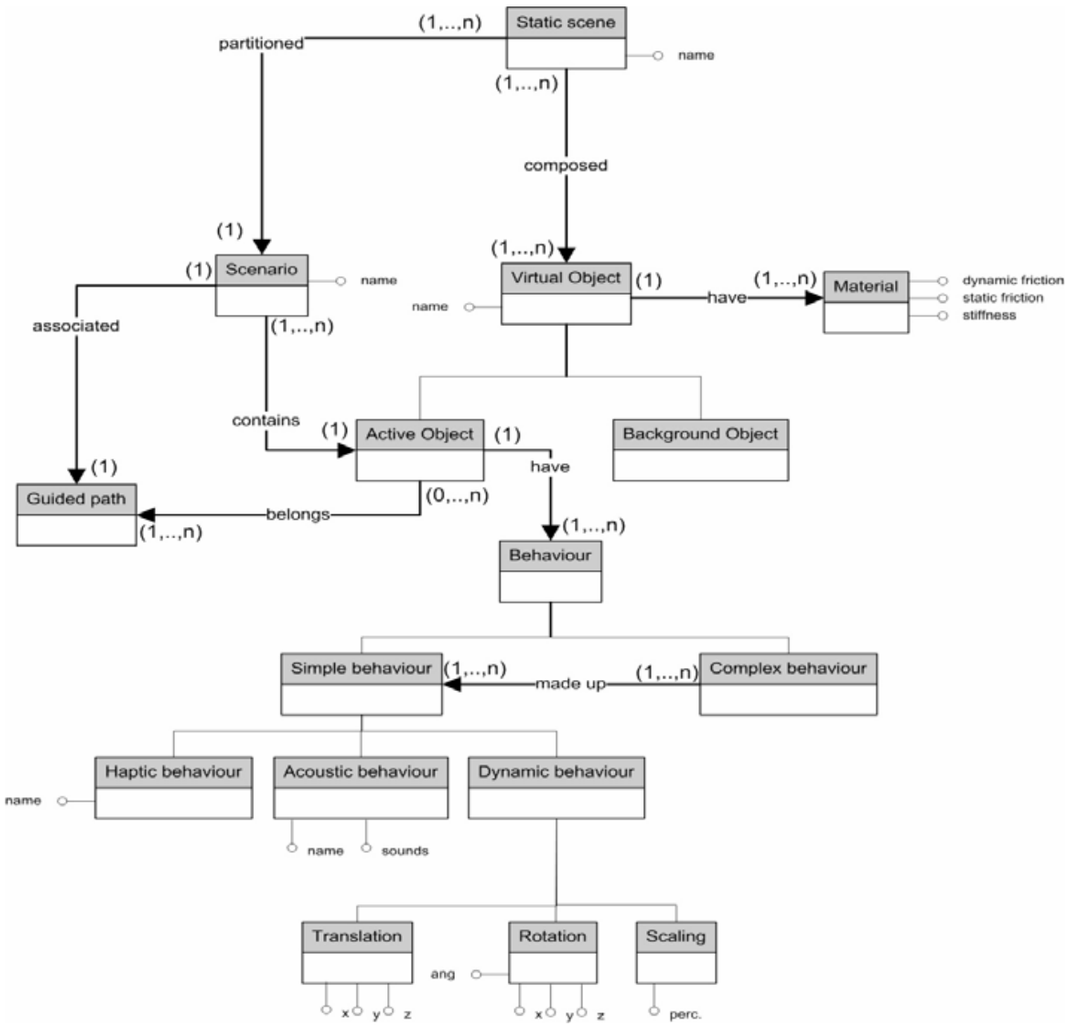


Fig. 4. The OMERO interaction conceptual model

effects such as vibration, attraction, viscosity or any given force field. Acoustic behaviors associate sounds and/or synthetic speech. Dynamics behaviors allow active objects to be manipulated in a strictly physical sense. Active Objects can act as targets in a route planning interaction belonging to a given Guided Path. In non-visual settings (section 3.1.3) targets are predefined by the designer to make easier to acquire particular information enclosed in the scene. Active Objects can be useful metaphors for designer and users. They give to designers an intuitive model to encapsulate selection and manipulation metaphors described in terms of their behavior.

4.2 The scene querying task

A rich in details virtual world generates a long sequence of local perceptions. Integrating all these data into a coherent and meaningful mental schema by means of touch, that does not provide a quick and global perception of the scene as sight does, is often a real challenge for blind. A VE can be seen as a database containing different kinds of information represented in the form of 3D objects shapes, together with their multimodal interface. Therefore, a user can explore a VE querying the scene to display different type of information. To better comprehend this task, we organize the scene using the Scenarios metaphor**.** Scenarios are sets of semantically related active objects that encapsulate informative contents. The user, at each specific time, can query the scene asking for a particular scenario and can focus his attention only on the information associated with it, temporarily discarding all the other data. This feature can reduce the discomfort that even seeing people may experience when faced with complex environments. Each Scenario of the model can be turned on and off (being touchable and visible or not). In this way, the scene can be tailored to focus on the data of interest having a progressive access to information.

5. Hapto/Acoustic interaction evaluation

Different applications of the OMERO system have been designed and proposed to visually impaired users to navigate VEs with the main aim of acquiring knowledge about particular domains, constructing their cognitive map and to exploit them, when it is possible, in navigating the corresponding real environments. The flexibility of digital models and multimodal interaction have been used to design VEs that allow blind to overcome some of the difficulties they encounter in acquiring spatial information in the every day life. Very different navigation contexts have been proposed to blind in several occasions. Their 3D virtual models, the related multimodal features and some details of the experimentation with blind people will be described in the following paragraphs.

5.1 Navigation task

The navigation of VEs in the applications we have designed for blind users is mainly based on the World in Miniature metaphor. Users first virtually navigate the most simple VE for a given context to construct the mental schema of the whole environment where then they can locate further objects. During the navigation they can interact with active objects, such as door or speaking walls. Experimental results validate the implemented approach.

The proposed approach is especially useful to navigate VEs reproducing sites characterized by large extension, complex topology and a huge amount of information to be conveyed to the user. This is the case of the visitable part of the ground floor of the Norman-Svevian

Castle located in Bari (Italy) whose virtual plant has been constructed from its detailed planimetry. The basic model includes: the entry area and the ticket office, the external and internal courtyards, the gallery of plaster casts, the chapel, the bathrooms and some connecting places (Figure 5a). Some doors can be opened and are defined as haptic-acoustic dynamic objects while doors to inaccessible environments are modelled as static objects with associated vocal explanatory messages. Transit areas without doors are modelled by bumps, defined as haptic-acoustic static objects.

Three sessions of tests have been accomplished on three different groups of users during two different events. The two groups tested during the first event were composed by four visually impaired people each, while the third group tested during the second event was composed by twelve blind. Any of them had previous knowledge of the castle.

The task they had to accomplish during the first two test sessions was to navigate the VE to acquire information on disposition and dimension of the castle environments and on the location of some objects. In the subsequent real visit they exploited knowledge coming from the virtual experience to consciously move in the castle. Models who they interacted with during the two test sessions were different. During the first test session, the model included most of the objects located inside the reproduced environments (trees, hedges, pots, ...) that were represented by simple solid shapes and defined as acoustic active objects (touching them, users automatically received vocal messages clarifying their identity). During the second one, the VE was modified removing any object different from doors and bumps and, after a first navigation of this model, users could explore zoomed models of some environments they required, with all the corresponding objects, to integrate the new pieces of information within the overall schema (Figure 5b).

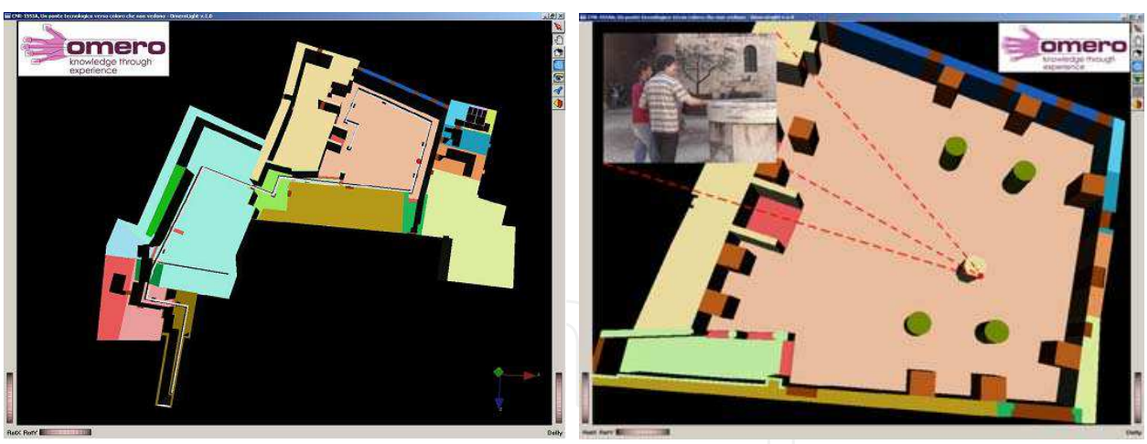


Fig. 5. (a) The 3D plant of a part of the Swebian castle (b) A blind finds the well in the internal courtyard of the real castle after the exploration of the corresponding zoomed VE.

All the people in the two groups were able to accomplish the proposed tasks in a satisfactory way, but those belonging to the first one had to overcome some difficulties due to the large number of objects placed in the model, the vocal messages frequently activated touching them and the small size of some environments of the castle with respect to the haptic device workspace. Users in the second group could better focus on shape and disposition of the rooms and on the haptic-acoustic interaction with doors and bumps

without being annoyed by the objects and the related vocal messages. The main strategy used to explore the VE was following its borders, in order to understand the related shape.
Both the test sessions with the multimodal system were followed by a real visit; visually impaired people could verify the mental schema of the environments they had constructed during the virtual experience. Each blind person accomplished the real visit assisted by the most congenial support (companion, cane or guide dog) to ensure its safety: in all cases, the blind was able to decide autonomously the path across the castle without any help. Users had some difficulties during the real exploration of the largest areas in the castle due to the fact that in this case it is not simple to recognize references and correctly match real and VEs in a short time. In any case, they outlined that all the implemented features (haptic/vocal synchronization, change of the scale of the model, insertion/removal of details, …) allowed them to construct an effective mental representation of the real environment and to navigate it in a conscious way. Some users stated to be interested in a new virtual visit, following the real one, to definitely assess their castle mental schema.
During the second test session, users were able to explore the virtual model in a faster and more effective way with respect to those of the first one; also their capacity to recognize the disposition of the environments in the real visits was enhanced. This could be probably due to the more intuitive and direct interaction with a simpler model, characterized by a much smaller amount of details and automatic vocal messages.
The task users had to accomplish during the third test session was to navigate the VE to acquire information not only about the topology of the castle but also about its history. Moreover users could virtually train for a little exhibition dedicated to blind people prepared in the gallery of plaster casts, with the aim of integrating information acquired from virtual and real visit. In this case, the VE presented a guided path (Figure 5a) that by means of 17 attractive target points placed along the predefined route, realized by means a sort of pipe, allowed blind to fast identify the salient portions of the castle. Whenever they remained trapped in a target they could activate a vocal message giving information on the place where they were. Then they could free navigate the virtual plant without objects inside enriched with respect the previous tests by means of on demand messages about dimensions of the environments, materials used for their construction and historical information. Finally they could experience the enlarged model of the gallery of plaster casts where some objects that they could explore during the real visit were located; their presence was highlighted by a suitable attraction effect that was noticeable whenever the avatar was into their neighbourhoods.
Users 1, 2 and 8 started their navigation moving too fast along the path and losing some targets but after few minutes they were able to move with the right speed. User 3, a guy with low vision, was able to reach each target by exploiting his residual sight. User 4 claimed that the experienced guided path should be integrated with vocal messages informing to the next direction to be taken. User 5 paid attention to ask for vocal messages whenever he reached a new target, and stated that this type of functionality can be really useful when exploring unknown environments. User 9 followed the path in a systematic way and he could feel every target; during the free exploration his movements followed the direction dictated by the guided path. He claimed that more experience with the VE could enhance the comprehension of the targets displacement. Users 10 and 12 had analogous results. Users, in spite of some difficulties due to their first approach with multimodal

interaction, judged the system to be very effective in making them able to move in unknown environment and acquiring different kinds of information that can be added as a useful complement to those provided by the real exploration.
Another example of guided path is the support function introduced in the Apulia model that pushes the avatar toward the nearest town when pressing the spacebar on the keyboard, helping the user to find the closest place of interest. This facility proved to be very useful and was intensively used especially during the first phases of the VE exploration.

5.2 Scene querying task

A virtual scene reproducing a complex environment can contain a big amount of information to be conveyed to users. Providing all the contents at the same time can overburden the blind user, making hard an effective navigation and comprehension of the VE. To simplify the user interaction with the virtual scene and organizing information on the basis of their semantic meaning, VEs can be designed exploiting the scenarios metaphor.
Based on this metaphor the virtual model of the Apulia region has been realized and tested with blind users. It has been constructed starting from GIS data to allow visually impaired users to acquire a proper knowledge of the territory from different points of view (Figure 6). It is multi-layered and, simply pressing the function keys on the keyboard, it is possible to navigate among four different scenarios each reproducing data related to a particular semantic view of the region, progressively building a structured and complete mental schema of the territory and its peculiarities.
A first scenario of the region model concerns the shape and the disposition of provinces, their borders and the borders between Apulia, the neighborhood regions and the sea. All the borders and the provincial areas have been defined as acoustic active objects. A second view reports the hydrographic network of the region: rivers and lakes have been respectively realized as canyon and ditches in which the avatar can fall and move to provide perceptions about their course and shape. All these objects have been defined as acoustic active objects associated with vocal messages telling their names. Another level includes the location of the major towns. The last view shows the main connections between towns. Roads are haptically represented as canyons connecting two towns and are acoustically active: on demand, they provide a vocal message about name, kind of the road and connected towns.
The described model has been proposed to twenty visually impaired users. Some of them did not have any previous knowledge of the region features. Users started their test exploring the first view of the model, to construct the cognitive map of the shape of the whole region and of each province, their borders, names and the relative position. Then they went through the next scenarios to increase their knowledge of the VE. The scene querying interaction modality helped them during the VE exploration and it was quite transparent; users were informed on the informative content of the scenario so they were not confused by comparing and disappearing objects in the different scenes but they were able to organize the information in an effective learning. Only a user complained loosing the reference points when changing scenario and suggested the possibility of merging scenarios information. Most of the users were able to correctly locate rivers, lakes and towns with respect to the regional and/or provincial territories and with respect to each other. Almost all the users judged the proposed interaction a valid and more flexible alternative to tactile maps and found the haptic-acoustic interaction really stimulating.



Fig. 6. The hydrographic network (a) and the main road scenarios (b) of the Apulia region

5.3 Selection/Manipulation task

While navigating VEs, users interacted with active objects such as doors, speaking walls or attractive objects. Doors, modeled as hapto-acoustic active objects, have been associated to a vibration effect that allows users to distinguish them from walls; they can be opened on demand by clicking the stylus button while the probe is in touch with them, listening at the same time to the opening/closing sound that makes more realistic the action in the virtual world. Few trials were usually sufficient to allow users to correctly interact with doors and to usefully exploit automatically triggered and on demand messages. Some users had problems in opening doors, due to the fact that sometimes the pressing movement can bring the probe from the modeled door down to the pavement, failing the opening command. Other users, after the door was opened, remained still without going in the other room.

Other type of active objects can be only acoustic: they tell the name of the touched object, the place where the avatar is in the VE or can act as informative points giving the user more structured information about geometrical formulas, definitions, materials, history… Vocal messages can be triggered on touch or on demand. Experiences with blind users suggested that vocal messages triggered on touch can be confusing if the activation is due to an accidental collision between avatar and active object. Sometimes, the pressing movement can influence the right placement of the avatar with respect to the active object to prevent from activating the message. Haptic active objects, exerting an attractive force onto the avatar whenever it comes close, were inserted in some rooms of the castle to highlight the presence of artefacts or were used to easily find towns in the Apulia model. In general the active objects interaction resulted effective for users that moving with slow movements were able to feel the attraction and to correctly interact with them listening to messages on the object nature. Users that moved fast in the VE often failed to feel the attraction and found objects only when accidentally collided with them. Some users outlined that objects located along the borders of the VE can act as effective reference points, while the most central objects are not important to this aim even if they allow a complete comprehension of the site. The most active users were really intrigued by finding them both in the virtual and when possible in the real visit.

Often blind users judged multimodal displays in different ways. For this reason, in (De Felice et al., 2009) a method for a fast multimodal authoring of VE has been proposed based on decoupling the geometric scene representation from its multimodal rendering. Designers

can easily change hapto-acoustic cues using a visual editor in order to fit users feedback. A didactical application conveying information on plane geometry has been realized to test this feature. It has been first tested on a visually impaired user that had never used haptic interfaces (PHANToM Omni) before. User feedback highlighted that haptic cues such as vibration and magnetic forces can be hard to appreciate if associated with tiny features such as the circle radius line as in (Yu & Brewster, 2002). The system facility was used to find in real time the configuration of haptic effects that allowed the user a better understanding of the geometrical VE. Moreover, synthetic speeches helped him in integrating information coming from the haptic feedback with higher level information on the related concepts.
To face with the problem of the limited physical workspace that constrains models dimension, Dragging and Scaling functionalities have been implemented as in (Magnusson & Rassmus-Gröhn, 2003). The dragging technique allows to make always available a part of any VE in the workspace for the virtual exploration. To simplify the perception of details that would be too small in a complete view of the environment and simplify the understanding of spatial data, the model can be presented at a larger scale changing the relative size of the models with respect to the avatar. Moreover, in addition to what implemented in (Magnusson & Rassmus-Gröhn, 2003), if the user requires the scaling while touching an object, the VE is scaled according to the position of this contact point, otherwise the scaling is made according to the position of the centre of the scene. This meaningful reference prevents the user from being confused by uncontrolled movements of the VE. A 10% scale factor has been applied for each zoom in/out step. Preliminary experimental tests with blind users showed no preferences between the two dragging techniques: the choice of the most comfortable method seemed to be influenced by subjective user characteristics. However, further studies are needed to better ground these conclusions.

6. Conclusion

In this chapter an overview of interaction techniques in the field of multimodal 3D VEs for non-visual interaction has been presented, based onto classic users tasks in VEs (navigation, selection and manipulation of virtual entities). For each of them and where possible a comparison between hapto-acoustic versions in visual and non-visual settings has been accomplished. This will facilitate the understanding of each technique and the way tasks normally carried out by vision can be completed with sensorial substitution.
Navigation tasks refer to strategies used to move in VEs. Direct camera manipulation techniques for blind or eye busy users should be avoided preferring an exocentric viewpoint with world in miniature metaphors. Guided paths and attractive targets techniques aid blind to compensate the lack of a global glance of the scene.
Selection task refers to the acquisition or identification of an object in the VE. Visual settings propose the integration with hapto-acoustic feedbacks such as bumps, magnetic lines, attractive forces, earcons and spatial sounds. Experimental results show that in visual settings hapto-acoustic cues do not provide significant benefits in improving the selection time and in disambiguating multiple selected and occluded targets. Users heavily rely on the visual feedback while the multimodal feedback helps to highlight the selection confirmation. In non-visual settings, selection is a quite passive procedure and attractive forces can be applied to objects in order to highlight their presence. To realise an active selection, a coarse haptic glance of the virtual scene is necessary. To this aim vibration tips

help locating objects for selection while attractive forces can guide the user to the selected object. Positional sounds highlight the object position and the confirmation of its selection.
Manipulation task allows to acquire the knowledge of selected objects though their translation, rotation and scaling or can refer to a more general concept of interaction. Few works exist for non-visual multimodal manipulation of virtual objects. In non-visual contexts, objects cannot be moved around but must remain fixed to define reference points. Egocentric viewpoint could be investigated to exploit the proprioceptive subsystem.
The most effective techniques for user tasks taken from literature and experimented in our past works have been chosen and integrated to build the OMERO multimodal interaction framework. The Scene Querying user task has been introduced for a gradual access to scene information based on the scenarios metaphor that decomposes the virtual world in sets of semantically related objects. Moreover, the Active Object metaphor has been developed to support selection and manipulation tasks allowing to encapsulate multimodal behaviours to better design selection and manipulation metaphors. Based on these features different applications have been developed and a large number of experiments with blind people have shown the effectiveness of the implemented interaction techniques for user tasks such as Navigation, Object Selection\Manipulation and Scene Querying. Virtual guided tours have proved to help blind users to acquire a first coarse schema of the salient areas of the VE while vocal messages provided useful cues to support way finding. Active Objects have been able to convey information in a effective and compact way. The scenarios mechanism has been well accepted and exploited for acquiring a better knowledge of Ves. Future works aim to develop scenarios mixing information to answer more complex user queries.

7. References

Akamatsu, M., MacKenzie, I.S. & Hasbrouc, T. (1995). A Comparison of Tactile, Auditory, and Visual Feedback in a Pointing Task using a Mouse-type Device, *Ergonomics,* Vol. 38, pp. 816-827

Anderson, T.G., (1998). Flight: An advanced human-computer interface and application development environment, *Master's thesis*, University of Washington, 1998

Bowmann, D.A. & Hodges, L.F. (1997). An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments, *Proceeding of the Symposium on Interactive 3D Graphics,* Providence, RI, USA, April 1997

Bowmann, D., Johnson, D. & Hodges, L. (1999). Testbed Evaluation of the VE Interaction Techniques, *Proceeding of the 1999 ACM Symposium on Virtual Reality Software and Technology*, ACM Press, pp. 26-33

Bowmann, D.; Kruijf, E., La Viola, J. & Poupyrev, I. (2001). An Introduction to 3-D User Interface Design. *Presence: Teleoperators and Virtual Environments*, Vol. 10, No. 1, pp. 96-108, February 2001

Bowmann, D.; Kruijf, E., La Viola, J. & Poupyrev, I. (2005). 3D User Interfaces: Theory and Practice. Addison-Wesley

Cockburn, A. & Brewster, S. (2005). Multimodal Feedback for the Acquisition of Small Targets, *Ergonomics,* Vol. 48, pp. 1129-1150

Conti, F., Barbagli, F., Morris, D. & Sewell, C. (2005). CHAI 3D: An Open-Source Library for the Rapid Development of Haptic Scenes, *IEEE Proc. WorldHaptics*, Italy, 2005.

Coomans, M.K.D. & Timmermans, H.J.P. (1997). Towards a Taxonomy of Virtual Reality User Interfaces, *Proc. 4th Int. Conf. on Information Visualisation*, England, Aug. 1997

De Boeck, J., Raymaekers, C. & Coninx, K. (2001). Expanding the haptic experience by using the phantom device to drive a camera metaphor, *Proceedings of the sixth PHANToM User Group Workshop*, Aspen, CO, USA, October 2001

De Boeck, J. & Coninx, K. (2002). Haptic camera manipulation: Extended the camera in hand metaphor, *Proceedings of EuroHaptics 2002*, pp. 36-40, Edinburgh, UK, July 2002

De Boeck, J., Raymaekers, C & Coninx, K. (2003). Blending speech and touch together to facilitate modelling interactions, *Proceedings of HCI International 2003,* Vol. 2, Crete, June 2003, pp. 621-625

De Boeck, J.; Cuppens, E., De Weyer, T., Raymaejers, C. & Coninx, K. (2004). Multisensory interaction metaphors with haptics and propioception in virtual environments, *Proceedings of NordiCHI 2004*, Tampere, FI, October 2004

De Boeck, J.; Raymaekers, C. & Coninx, K. (2005). Are Existing Metaphors in Virtual Environments Suitable for Haptic Interaction, *Proceedings of the 7th International Conference on Virtual Reality*, pp. 261-268

De Felice, F.; Gramegna, T., Renna, F., Attolico, G. & Distante, A. (2005). A Portable System to Build 3D Models of Culturale Heritage and to Allow Their Explorations by Blind People, *IEEE Proc. of HAVE 2005*, Ottawa, Ontario, Canada, October 2005.

De Felice, F.; Renna, F., Attolico, G. & Distante, A. (2007). A haptic/acoustic application to allow blind the access to spatial information, *Proceeding of WorldHaptics 2007*, Tzukuba, Japan, March 2007

De Felice, F., Attolico, G. & Distante, A. (2009). Configurable Design of Multimodal Non-Visual Interfaces for 3D VEs, *Proceedings of HAID 2009 4th International Conference on Haptic and Audio Interaction Design,* Dresden, Germany, September 2009, pp. 71-80

Esposito, C. (1996). User interfaces for virtual reality systems, *Human Factors in Computing Systems, CHI96 Conference Tutorial Notes*, April 1996

Forsberg, A., Herndon, K & Zeleznik, R. (1996). Aperture Based Selection for Immersive Virtual Environments, *Proceeding of the 1996 ACM Symposium on User Interface Software and Technology,* ACM Press, pp. 95-96

Gabbard, J. & Hix, D., (1997). A Taxonomy of Usability Characteristics in Virtual Environments, Virginia Polytechnique Institute and State University, Nov. 1997

Garre, C. & Otadui, M.A. (2009). Toward Haptic Rendering of Full-Hand Touch, *in Proceedings of CEIG'09*, San Sebastian, September 2009

H3DAPI: www.h3d.org, accessed the 5 Agoust 2010.

HaptiMap European project, www.haptimap.org, accessed the 5 Agoust 2010

Hinkley, K., Pausch, R., Goble, J. & Kassell (1994). Passive Real-World Interfaces Props for Neurosurgical Visualization, *Proceeding of the 1994 ACM Symposium on Human Factors in Computing Systems,* ACM Press, pp. 452-458

Höllerer, T., Feiner, S., Terauchi, T., Rashid, G. & Hallaway, D. (1999). Exploring MARS : Developping Indoor and Outdoor User Interfaces to a Mobile Augmented Reality System, *Computer and Graphics*, Vol. 23, No. 6, 1999, pp. 779-785

Igarashi, T., Kadobayahi, R., Mase, K. & Tanaka, H. (1998). Path Drawing for 3D Walktrough, *Proceeding of the 1998 ACM Symposium on User Interface Software and Technology*, ACM Press, pp. 173-174

Immersion: www.immersion.com, accessed the 5 August 2010.

Jacobson, R.D. (2002). Representing spatial information through multimodal interfaces, *Proceedings of Sixth International Conference on Information Visualisation*, 2002

Jaimes, A. & Sebe, N. (2007). Multimodal human-computer interaction: A survey, *Computer vision and image understanding*, Vol. 108, No. 1-2, pp. 116-134

Klatzky, R. & Lederman, S. (1995). Identifying objects from a haptic glance, *Perception and Psychophysics*, 57, 1995, pp. 1111-1123

Kolcárek, P., Sochor, J. (2005). Velocity Driven Haptic Rendering, *Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, Dunedin, New Zeland, 2005.

Lahav, O. & Mioduser, D. (2008). Haptic-feedback support for cognitive mapping of unknown spaces by people who are blind, *Int. J. of Human-Computer Studies,* 23-25

Lakoff, G. & Johnson, M. (1980). Metaphors we live by, The Chicago University Press

Lecuyer, A., Mobuchon, P., Megard, C., Perret, J., Andriot, C. & Colinot, J. (2003). Homere: A Multimodel System for Visually Impaired People to Explore Virtual Environments, *Proceedings IEEE Virtual Reality,* pp. 251-258, 2003.

Liang, J. & Green, M. (1994). JDCAD: A Highly Interactive 3D Modeling System, *Computer and Graphics,* Vol. 18, No. 4, pp. 499-506

Magnusson, C. & Rassmus-Gron, K. (2003). Non-visual zoom and scrolling operations in a virtual haptic environment, *in Proc. of EuroHaptics 2003*, Dublin, Ireland, July 2003.

Magnusson, C. & Rassmus-Gron, K. (2004). A Dynamic Haptic-Audio Traffic Environment, *Proceedings of Eurohaptics 2004*, Munich, Germany, June 2004

Magnusson, C. & Rassmus-Gron, K. (2005). Audio haptic tools for navigation in non visual environments, *Proceedings of Enactive05, 2nd International Conference on Enactive Interfaces*, Genoa, Italy, November 2005.

Menelas, B., Picinali, L., Katz, B. & Bourdot, P. (2010). Audio Haptic Feedbacks in a Task of Targets Acquisition, *Proceedings of IEEE Symposium on 3D User Interface*, Waltham, MA, USA, March 2010, pp. 51-54

Mine, M. (1995). ISAAC: A virtual environment tool for the interactive construction of virtual worlds, *Technical Report TR95-020, UNC Chapel Hill Computer Science*

Mine, M.R. (1996). Working in Virtual World: Interaction Techniques Used in the Chapel Hill Immersive Modelling Program, *Technical Report TR96-029,* August 1996

Mine, M.R. & Brooks, F.P. (1997). Moving objects in space: Exploiting proprioception in virtual environment interaction, *Proocedings of the SIGGRAPH 1997 annual conference on Computer Graphics,* Los Angeles, CA, USA, August 1997

Moustakas, K.; Nikolakis, G., Kostopoulos, K., Tzovaras, D. & Strinzis, M. (2007). Haptic Rendering of Visual Data for the Visually Impaired. *IEEE MultiMedia,* 1, pp. 62-72

Murai, Y.; Tatsumi, H., Nagai, N. & Miyakawa, M. (2006). A Haptic Interface for an Indoor-Walk-Guide Simulator, *Proc. of ICCHP 2006,* Lintz, Austria, July 2006, Springer

Nesbitt, K.V. (2005). A Framework to Support the Designers of Haptic, Visual and Auditory Displays, *In Carter and Fourney {3}*, pp. 54-64

Okamura, A. & Cutkosky, M. (1999). Haptic exploration of fine surface features, *Proceedings of IEEE Int. Conf. on Robotics and Automation*, pp. 2930-2936, 1999

Ott, R., Vexo, F. & Thalmann, D. (2010). Two-handed Haptic Manipulation for CAD and VR Applications, *Computer-Aided Design & Applications,* Vol. 7, No. 1, pp. 125-138

Oviatt, S.L. (1999). Ten Myths of Multimodal Interaction, *Comm. ACM*, Nov. 1999, pp. 74-81.

Panëels, S., Roberts, J.C. & Rotgers, P.J. (2010). Review of Design for Haptics Data Visualization, *IEEE Transactions on Haptics*, Vol.3, No. 2, pp. 119-137

Parente, P. & Bishop, G. (2003). BATS: The Blind Audio Tactile Mapping System, *Proceedings of the ACM Southeast Regional Conference,* March 2003, ACM

Piateski, E. & Jones, L. (2005). Vibrotactile Pattern Recognition on the Arm and Torso, *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems,* pp. 90-95.

Pierce, J., Forsberg, A., Conway, M., Hong, S., Zeleznik, R. & Mine, M. (1997). Image Plane Interaction Techniques in 3D Immersive Environments, *Proceeding of the 1997 ACM Symposium on Interactive 3D Graphics,* ACM Press, pp. 39-44

Pierce, J., Stearns, B. & Paush, R., (1999). Voodoo Dolls: Seamless Interaction at the Multiple Scales in Virtual Environments. *Proceeding of the 1999 ACM Symposium on Interactive 3D Graphics,* ACM Press, pp. 141-145

Pokluda, L. & Sochor, J. (2005). Spatial Orientation in Buildings Using Models with Haptic Feedback, *Proceedings of WorldHaptics 2005*, Pisa, Italy, 2005.

Poupyrev, I., Billinghurst, M., Weghorst, S. & Ichikawa, T. (1996). The Go-Go Interaction Technique: Non Linear Mapping for Direct Manipulation in VR, *Proc. of the 1996 ACM Symposium on User Interface Software and Technology,* ACM Press, pp. 79-80

Poupyrev, I., Tomokazu, N. & Weghorst, S. (1998). Virtual Notepad: Handwriting in Immersive VR, *Proceeding of the 1998 IEEE Virtual Reality Annual International Symposium,* IEEE Press, pp. 126-132

Prada, R. & Payandeh, S. (2005). A Study on Design and Analisys of Virtual Fixtures for Cutting in Training Environments, *Proc. of Worldhaptics 2005,* Pisa, Italy, pp. 375-380

Roberts, J.C. & Panëels, S. (2007). Where are we with Haptic Visualization?, *Proceedings of WorldHaptics 2007,* Tzukuba, Japan, March 2007, pp. 316-321

Salisbury, K.; Conti, F., & Barbagli, F. (2004). Haptic Rendering: Introductory Concepts, *IEEE Computer Graphics and Applications*, Vol. 2, 2004, pp. 24-32

Sensable: www.sensable.com, accessed the 5 August 2010

Sjöström, C. (1999) The IT Potential of Haptics – touch access for people with disabilities. Licentiate thesis, available at http://www.certec.lth.se/doc/touchaccess

Sherman, W.R. & Alan, B.C. (2003). Understanding Virtual Reality, Interface, Application, and Design, Morgan Kaufmann

Tzafestas, C.S. (2003). Whole-Hand Kinesthetic Feedback and Haptic Perception in Dextrous Virtual Manipulation, *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans,* Vol. 33, No. 1, January 2003

Vanacken, L., Grossman, T. & Coninx , K. (2009). Multimodal Selection Techniques for Dense and Occluded 3D Virtual Environments, *International Journal of Human-Computer Studies,* Vol. 67, 2009, pp. 237-255

Vidholm, E., Tizon, X., Nyström, I. & Bengtsson, E. (2004). Haptic Guided Seeding of MRA Images for Semi-Automatic Segmentation, *in ISBI*, pp. 288-291

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 Acquisition Task, *Proceedings of Eurohaptics 2002,* pp. 23-28

Ware, C. & Osborne, S., (1990). Exploration and Virtual Camera Control in Virtual Three-Dimensional Environments, *Computer Graphics, Vol. 24, No. 2.*

Welch, G., Vicci, L., Brumback, S., Keller, K. & Colucci, D. (2001). High-performance Wide-Area Optical Tracking: The HiBall Tracking System, *Presence: Teleoperators and Virtual Environments,* Vol. 10, No. 1, pp. 1-21

Yu, W. & Brewster, S. (2002). Multimodal Virtual Reality Versus Printed Medium in Visualization for Blind People, *Int. ACM Conf. on Assistive Technologies*, pp. 57-64

Zhai, S., Buxton, W. & Milgram, P. (1994). The "Silk Cursor": Investigating Transparency for 3D Target Acquisition, *Proceedings of the 1994 ACM Conference on Human Factors in Computing Systems,* ACM Press, pp. 459-464

Zilles, C.B. & Salisbury, J.K. (1995). A constraint-based god-object method for haptic display, *Proceedings of the International Conference on Intelligent Robots and Systems*, Vol. 3

IntechOpen



Virtual Reality

Edited by Prof. Jae-Jin Kim

ISBN 978-953-307-518-1
Hard cover, 684 pages
Publisher InTech
Published online 08, January, 2011
Published in print edition January, 2011

Technological advancement in graphics and other human motion tracking hardware has promoted pushing "virtual reality" closer to "reality" and thus usage of virtual reality has been extended to various fields. The most typical fields for the application of virtual reality are medicine and engineering. The reviews in this book describe the latest virtual reality-related knowledge in these two fields such as: advanced human-computer interaction and virtual reality technologies, evaluation tools for cognition and behavior, medical and surgical treatment, neuroscience and neuro-rehabilitation, assistant tools for overcoming mental illnesses, educational and industrial uses In addition, the considerations for virtual worlds in human society are discussed. This book will serve as a state-of-the-art resource for researchers who are interested in developing a beneficial technology for human society.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Fabio De Felice, Floriana Renna, Giovanni Attolico and Arcangelo Distante (2011). Hapto-Acoustic Interaction Metaphors in 3D Virtual Environments for Non-Visual Settings, Virtual Reality, Prof. Jae-Jin Kim (Ed.), ISBN: 978-953-307-518-1, InTech, Available from: http://www.intechopen.com/books/virtual-reality/hapto-acoustic-interaction-metaphors-in-3d-virtual-environments-for-non-visual-settings

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen