Navigation help in 3D Worlds:
some empirical evidences on use of sound

C. Ardito, M. F. Costabile, A. De Angeli, F. Pittarello

Abstract: The concept of Interaction Locus (IL) has been introduced to help the users to orient, navigate, and identify relevant interaction areas in 3D Virtual Environments (VE). The IL is a multimodal concept: it adds to the 3D visual scene parallel information channels that are perceived by other senses. In particular, the IL emphasizes the role of music as a navigation aid in a VE. This paper reports three user-evaluations of different IL enriched virtual worlds, and in particular of the role of the IL auditory component. Results suggest that audio in 3D plays not only an aesthetic role, which the users greatly appreciate, but also a functional role simplifying navigation and helping the users to recognise scenes in the environment. Such a functional role however is subordinated to a proper understanding of the semantic link between music and virtual space. While these experiments refer to desktop virtual reality environments, their findings are general enough to inform the design of navigational tools for other segments of the mixed reality domain.

Keywords: Auditory Interfaces, Experimental Evaluations, Interaction Locus, Usability, Virtual Environments.

1. Introduction

In recent years, the number of cultural institutions offering virtual tours inside 3D museums is greatly increased. Interesting examples on the Web are the *Louvre Museum* (in Paris) and the *Van Gogh Museum* (in Amsterdam). Virtual worlds, however, are still far from being usable and lack of orientation is a major problem [6]. Design and usability guidelines for 3D Virtual Environments (VEs) are in their infancy and little knowledge can be directly transferred from traditional 2D interfaces. Beyond incorporating traditional usability considerations, 3D VEs open a number of specific issues, such as the design of navigation techniques to support user's movements, the design of specific interaction techniques to support object selection and manipulation, as well as ad-hoc visual, auditory and haptic feedback. Usable VE interfaces need to enhance presence, immersion, and comfort, while minimizing confusion.

VE browsers often do not have a built-in standardized support for user orientation; this is implemented by VE designers and normally consists in visual signs placed in the scene. Analogously to what happens in the real world, designers create the environment using geometric elements, which are composed together to define different areas of the scene. In spite of these composition efforts, visitors of VEs often experience feelings of being "lost" when trying to figure out how to navigate back to where they might have been just a few links before [7], [5], [21]. Several factors hamper orientation in VEs [17], [18]. The most common include shortage or invisibility of buildings and natural elements, and the serialization of specific objects (e.g., the iteration of the same architectural typology as happens in some British suburbs).

Navigation requires the ability to maintain knowledge of one's location and orientation while moving throughout a space. It involves cognitive activities (e.g., route planning), rather than simply manual manoeuvring. In current VEs, users tend to devote much of their cognitive resources trying to figure out the spatial layout, to the detriment of their main task (e.g., acquiring information), as insufficient or inappropriate information is provided about the spatial structure of the environment or about the identity and location of objects [15]. Specific help has to be developed to support navigation, assisting users in acquiring knowledge of the spatial relations, which characterise the world, recognising relevant areas, and building a mental map of the environment. The Interaction Locus (IL) approach proposes a solution to this

problem associating a set of hypertextual, auditory and tactile components to specific areas of a 3D environment [17], [18]. These information-enhanced areas have the potential for enhancing the comprehension of the structure of a 3D scene and therefore facilitating navigation.

This paper reports three user evaluations of the IL approach and, in particular, of the role of the auditory component in supporting orientation and navigation in 3D environment. The paper has the following organization. Subsection 1.1 considers related works, addressing the issue of navigation in 3D virtual worlds and the role of sound in interaction; Section 2 describes the IL concept; Section 3 and 4 report the user evaluations of three different IL enriched worlds. Section 5 provides a final discussion and conclusions, indicating also directions for future work.

1.1 Auditory interfaces

The idea of using audio to improve interaction in computer-supported tasks is not new [4], more original is the extension of this idea to 3D VEs. Different types of technologies (e.g., speech recognition, 3D auditory localization and synthetic speech output) are available to provide aural feedback to users [14]. Aural feedback includes sounds resulting from users' own actions, others' actions, and natural or ambient sounds [9]. Sound is used to enhance visual perception or as an effective sensory substitution, when no tactile or visual feedback is available [2]. In a VE, sound can be used as an output medium in two primary ways, localization and sonification [2]. Localization refers to generating 3D sound to give feedback about the presence of another user in a collaborative VE (i.e. utilizing walking sounds) or the position of an interesting object identifying it with a special acoustic mark. Sonification instead refers to turning certain types of information into sounds.

One major question to be answered when designing auditory interfaces is: what sounds should be used? Brewster et al. in [3] outlines some available options. Gaver's auditory icons are based on environmental sounds that have a semantic link with the object they represent. These icons have been used in several systems, such as the SonicFinder [10], SharedARK [11], ARKola [12] and are an effective form of presenting information. Earcons are another important way to provide audio feedback [1], [22], [23]. Blattner et al. in [1] define earcons as "non-verbal audio messages that are used in the computer/user interface to provide information to the user about some computer objects, operations or interaction". Earcons are composed of motives, which are short, rhythmic sequences of pitches with variable intensity, timbre and register. They are abstract, synthetic tones that can be used in structured combinations to create sound messages representing parts of an interface.

Earcons provide a powerful sonification method. They are used for adding sound to both data and interfaces and can be combined to produce complex audio messages. Related items are given related sounds and hierarchies of information are represented. Basic earcons for a set of simple operations, such as 'open', 'close', 'file' and 'program', can be created and combined to produce, for example, earcons for 'open file' or 'close program'. In this way, complex messages made of sub-units are built up.

2. Interaction Locus

The Interaction Locus concept has been introduced for enhancing navigation in 3D environments. It helps users to recognize relevant areas in a VE, allowing them to build a mental map of the environment for navigation purposes [17], [20]. Each area characterised by a specific morphology and interaction activity is associated with a set of hyperlinked components, such as text, visual, auditory and tactile information. The resulting information-enhanced is called Interaction Locus. The hyperlinked text provides a label and additional information about the location, whereas the auditory, visual and tactile components are symbolic marks of the location. ILs are organized in sets. Typically, a single 3D scene is mapped with a number of ILs that can be visited without any specific order or following a particular sequence, according to the nature of the interaction [20]. The resulting set of ILs becomes the key element for navigation inside 3D environments.

The auditory component, rather than being a simple aesthetic stream, is designed to help the users to build a cognitive map of the environment. In this way, the IL approach extends the use of earcons for orientation in 3D scenes. Audio acts as an additional mark-up of the place the user is visiting. The approach is more sophisticated than a multimedia one, because sound relates to the semantic characterization of the space. Therefore, the IL approach introduces multimodality in 3D environments, adding to the visual scene parallel information perceived by other senses.

The IL concept originated in the realm of the desktop virtual reality and was applied to cultural heritage [19]. Significant applications were implemented for the Palazzo Grassi web site, a famous museum in Venice and their evaluation is reported in this paper. In spite of its original paradigm and application area, the IL approach can be extended to different segments of the mixed reality domain. In particular, Fogli and colleagues [8] have applied the concept to the so-called lightweight mixed-reality (i.e., users move in a real environment while interacting with small location-aware devices such as PDAs) and have presented a case study related to archaeology sites. Additional application areas are currently under investigation, including tourism and e-commerce. The IL approach is also the base of a unified methodology allowing content experts to author hypermedia-enhanced 3D worlds for different paradigms of the mixed reality [20].

The following sections report three user evaluations of the IL approach and of the auditory component in supporting orientation and navigation in VE.

3. Einstein's Tower evaluation

The user studies reported in this paper aimed to evaluate the concept of IL as a navigation aid in different 3D environments. The first study was an exploratory investigation of the effect of audio on navigation and content fruition. The *Einstein Tower,* a VRML counterpart of the sun observatory built in Potsdam by Erich Mendelsohn (Figure 1), was used as a test bed. It was designed in 1997 for the "German Expressionism" exhibition at Palazzo Grassi [16]. A selection of works from the real exhibition is displayed in the research rooms and laboratories of the tower. For each exhibit, the visitors can ask for information, clicking with the mouse on the corresponding image. Moreover, words link to other information, in the usual way provided by hypertextual tools.
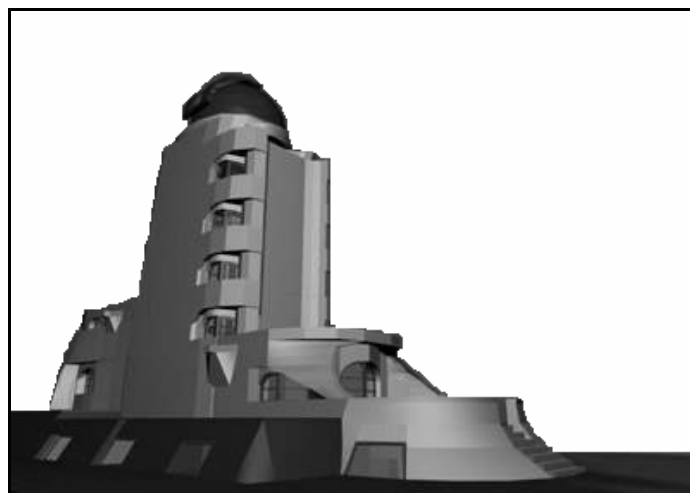

Figure 1: Einstein Tower

A set of ILs is available in the *Einstein Tower* world. Entrance in a locus is notified with a looped sound and a text appearing on the left of the 3D scene (Figure 2). The audio component is implemented in the form of a background classic music piece univocally associated to a Locus and automatically activated each time the visitor enters it without a fade-in and fade-out effect. The music played in the Locus has the same loudness independent of the visitor's position in the scene. When the visitor leaves the IL to enter another 3D scene not characterised by any interaction activity (i.e. going along a corridor or a staircase) the

music goes off suddenly. Different fragments of musical compositions related to the expressionist period characterise specific interaction areas.

The interface provides several visit styles (Figure 2). An automatic tour option drives the users through the rooms of the tower. Besides, the users can visit the exhibition at their own pace, going back and forth a predefined path, by clicking on the arrows placed below the virtual scene. They can use a free navigation mode too - using the arrow keys of their keyboard - or they can activate the building map and click over it to reach a specific location.
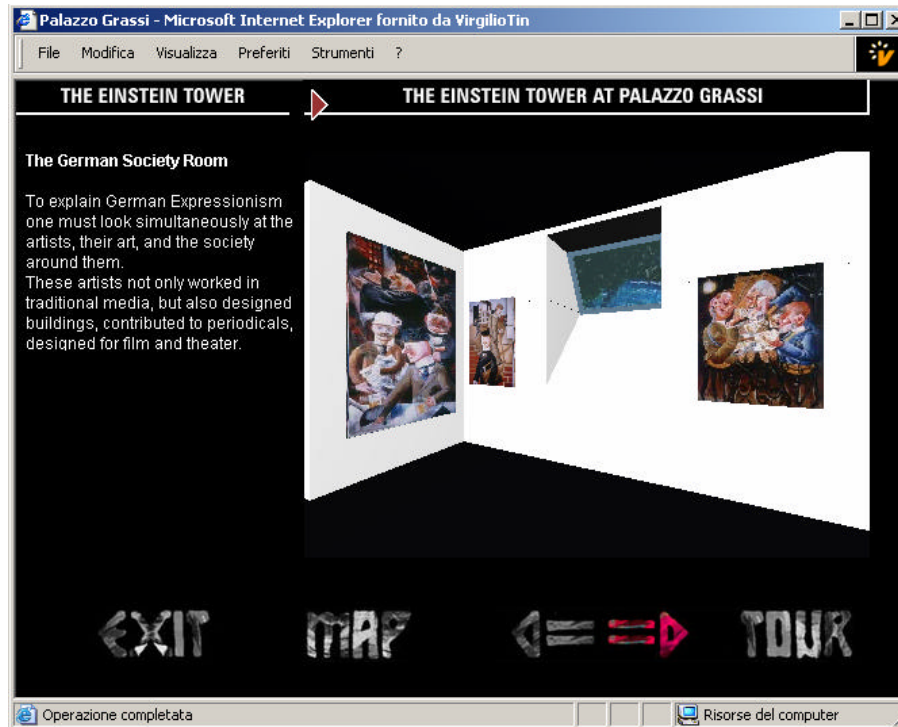

Figure 2: A scene of the Einstein Tower world

3.1   Experiment 1

The first experiment was aimed at investigate what (if any) effect the music had on the user behaviour and on the fruition of the *Einstein Tower* world. In theory, the music can have three (if any) positive effects on the user experience and behaviour. Namely, it can play a functional role, an aesthetic role or a combination of them.

In case of a *functional effect*, music should act as a major cue in supporting orientation, navigation and identification of relevant interaction areas. This hypothesis is at the basis of the IL approach. It assumes that audio can be semantically associated with space, and that the user is capable of recognising and exploiting this association. In case of an *aesthetic effect*, music should act as a pure entertainment channel. Thus, it should not affect performance but improve the overall experience of the Einstein Tower. These two effects are not exclusive. On the contrary, in a properly designed virtual world they are likely to co-occur. On the other side, there is the possibility that music could hamper navigation. Indeed, as navigation in VEs is difficult, an increase in stimulation could augment the user's mental workload and compromise performance.

To test these hypotheses, three different implementations of the *Einstein Tower* were designed and compared (Table 1). System A is a unimodal version; no audio is provided. System B is the original IL enriched multimodal implementation. System C is a multimedia version in which the same musical piece is played as a background accompaniment throughout the navigation.

Table 1: Experimental conditions in Experiment 1

Version	Audio	Music role
System A	Absent	/
System B	Present	Functional
System C	Present	Aesthetic

3.1.1 Method

Participants. Since the application topic was a little sophisticated, we selected on purpose participants who could get interested and motivated by it. They were 40 master and PhD students of Modern Art and History at the University of Bari, (16 M, 24 F).

Procedure. The experiment was composed of a training phase and two experimental sessions. During the training phase, participants used the guided tour option to visit all rooms and exhibits of the virtual museum in a sequence, without any control over the navigation. In this way participants could concentrate their attention to the exhibits in the virtual environment. The training was run in the same condition (no audio, functional music or aesthetic music) as the first experimental session; the music was presented to participants through headphones. The experimental sessions consisted of two sets of four equivalent tasks communicated by written instruction. Starting from the tower entrance, participants had to navigate in the tower for collecting information, such as rooms and paintings names, and to record it on an experiment booklet. To discover such information, participants had to locate the object and click on it. At the end, they filled in a user satisfaction questionnaire, requiring a general evaluation of the site with particular attention to ease of navigation, perceived sense of direction, and music quality. Two evaluators assisted in the experiment. The interaction was videotaped and data automatically logged.

Design. Three systems were tested (Table 1). System A (Unimodal system with no music) was used as a control condition to assess baseline values of the web site usability. All participants (N=40) used it together with one of the two audio/present systems and were randomly assigned to one of the two Music Role (Functional vs. Aesthetic) conditions. Half of the sample interacted with the IL enriched system playing functional music. The other half used the multimedia system, playing aesthetic music. Thus, Music Role was manipulated between-subjects. Audio (Present vs. Absent) was manipulated within-subjects. This design had the advantage to test several hypotheses simultaneously, minimising the number of participants, increasing the power of statistical comparisons, and controlling individual variability. To avoid carry-over effect, Audio test order was counterbalanced across participants and task set order was counterbalanced across experimental conditions.

3.1.2 Results

The evaluation addressed navigation performance and user satisfaction. The following variables were used to evaluate performance.

1. *Errors*: average number of wrong answers per task.
2. *Navigation speed*: average time spent navigating through the virtual path per task.
3. *Navigation steps*: average number of clicks on the arrows to go forward and backward through the virtual path per task.

Wrong answers (N=38) accounted for some 12% of the experimental questions, 18 occurred when participants navigated without music and 20 when they navigated with the music. Errors were homogeneously distributed among experimental conditions with no effect of music role.

Navigation speed was normalised by a logarithmic transformation and analysed by an analysis of variance with Audio (Present-Absent) as the within-subjects factor and Music Role (Functional-Aesthetic) as the between-subjects one. No significant effects emerged. Mean values are illustrated in Figure 3. Another analysis of variance was performed on navigation steps, yielding once again no significant effect.
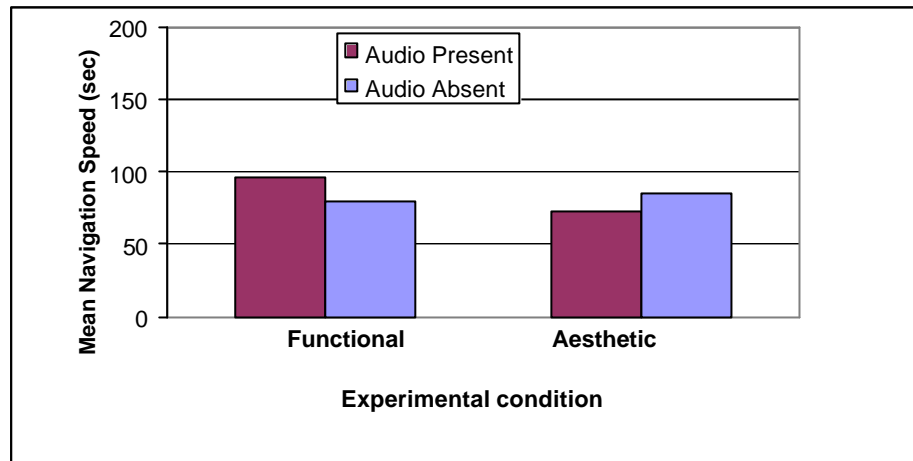
Figure 3: Navigation speed as a function of experimental conditions in Experiment 1

Questionnaires and interviews assessed user satisfaction. Some 93% of the sample explicitly declared preferring the audio version to the unimodal one, with no difference due to the role played by music. Similarly, no difference emerged with respect to the user satisfaction with the site and to the perception of the navigation task.

3.1.3 Discussion

This study did not support the IL hypothesis stating that audio plays a functional role and provides a support to user navigation. No differences in performance were found when participants navigated with or without music, nor between different types of musical accompaniments designed to play a functional or an aesthetic role. However, participants always declared to enjoy music independent of its intended role and to prefer the audio version to the unimodal one. Overall, results of this experiment supported the aesthetic hypothesis, claiming that a musical accompaniment in a 3D world increases the user satisfaction of it. Participants preferred the musical version, but music did not support navigation or exploration of the environment.

A possible explanation of these findings is that users did not understand the music/environment relationship as implemented in the *Einstein Tower*. In effect, music was quite sophisticated, and appreciation of the semantic link between musical pieces and physical environments was difficult. This difficulty may have hampered the recognition of the connection between music and space, thus decreasing the expected advantage of the IL approach. This idea was supported by user comments and observations made during the experiment. It appeared that users assigned to the IL conditions were sometimes distracted and annoyed by the music going on or off at some places during the navigation. This phenomenon was not associated to a causal trigger (i.e., entrance in the IL area), but it was perceived as random. To test the hypothesis that the negative results of Experiment 1 were due to a lack of understanding of the intended value of sound in IL, a new evaluation was run. This time participants were informed about the intended role of music, and invited to pay attention to the environment where different pieces of music appeared.

3.2 Experiment 2

Ten participants were recruited among graduate students at the department of Modern Art and History of the University of Bari. The procedure was the same as that of Experiment 1, but this time participants were told about the relationship between music and scene. Furthermore, they used only system A (no Music) and system B (Functional Music).

To evaluate the system in the most critical condition, all participants in the second evaluation used system B first. They were then compared with participants who followed the same test order in Experiment

1. The following results are based on 20 participants, 10 from each evaluation. For the sake of simplicity, we will call these two groups as *Informed* and *Uninformed* participants.

In the Informed condition, participants were informed about the functional role of music (Experiment 2). In the Uninformed condition, no explanation at all was given about the link between music and spaces (Experiment 1). The experimental hypothesis predicted that informed participants should be capable of exploiting the functional role of audio, thus being more efficient during navigation.

3.2.1 Results

To control individual variability in navigation, we initially compared performance without the music between the two groups of users, as regards number of errors, navigation speed and steps. No differences emerged, leading to the conclusion that participants' skills were comparable between the two groups.

Navigation speed and navigation steps were analysed by two independent sample t-tests. The distribution of navigation speed was normalised by a logarithmic transformation. Results were significant for both comparisons, $t_{(18)}= 2.22$ $p <.05$ for navigation speed and $t_{(18)}=1.9$ $p <.05$ (one tail) for navigation steps. Participants, who were informed about the role of music, executed their tasks faster than uninformed participants did (see Figure 4). Navigation in the informed sample was also more precise, as participants needed in average 11 steps less per task for reaching their objectives. Being informed of the role of the music also appeared to help participants to recognise specific loci, as they tended to provide less wrong answers than in the previous experiment (1 vs. 6).
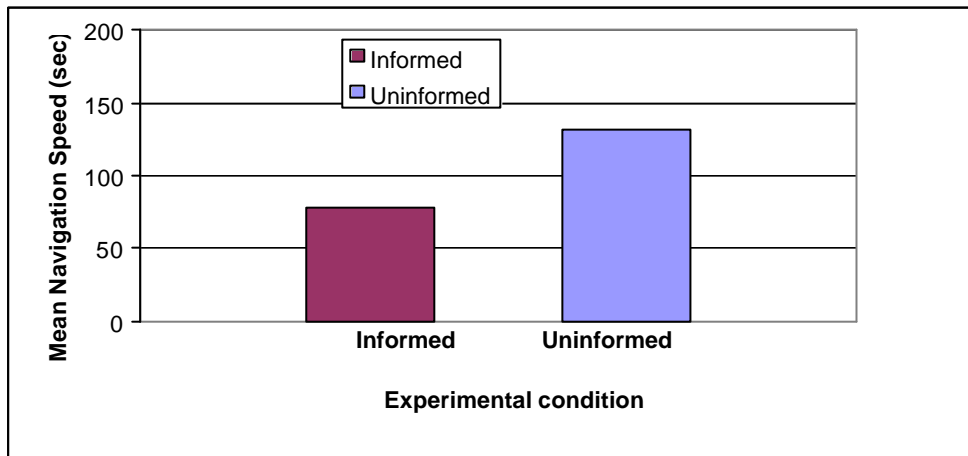

Figure 4: Navigation time as a function of experimental conditions in Experiment 2

3.2.2 Discussion

The experiment suggests that the IL concept has the potential for improving navigation in 3D environments. User performance was significantly improved by informing users of the music meaning. Informed participants were faster, probably because they did not waste time trying to understand what triggered the music on or off during the navigation. Furthermore, they were more precise in navigation, as reflected by the lower number of clicks on the navigation arrows, and by better accuracy in their answers.

These results suggest that the functional hypothesis on which the IL concept is based may be valid and that music could actually help people to navigate in 3D worlds. They also stress a basic problem of the *Einstein Tower* implementation, where the semantics of the music was too difficult. To gain a better understanding of the usability of the IL audio component, a new user study was run evaluating different 3D Virtual exhibitions: The *Chichen Itza* and *Newton's Cenotaph*.

4. Chichen Itza and Newton's Cenotaph evaluations

These studies addressed the role of the IL audio component in improving recognition of previously visited environments, a fundamental component of navigation and interaction. Slightly modified versions of two virtual exhibitions were used, namely *Newton's Cenotaph* and *Chichen Itza*. These 3D environments are part of the official web site of Palazzo Grassi [16]. The *Newton's Cenotaph*, conceived by I.E. Boullée as a tribute to the famous scientist, was modelled for the "Cosmos" exhibition and used as the scenario for a virtual art exhibition (Figure 5). A set of ILs identifies relevant areas. The users can choose between two visit styles: the automatic tour option, which allows them to visit the environment without additional interaction, or the guided tour paradigm, for controlling each step of navigation along a predefined path.

Figure 5: Newton's Cenotaph

The *Chichen Itza* environment (Figure 6), designed for the "Maya" exhibition, allows the visitor to walk through the buildings of an ancient Maya town. A set of ILs is superimposed to the city, identifying areas with a specific morphologic identity (e.g., the temples, the market, the pelota courtyards, etc.). While users are free to move without constraints, a predefined path for leading them to significant parts of the environment is available as a complementary navigation opportunity. In both environments, ILs are characterized by a spatial location, auditory and hypertextual components. ILs in the *Chichen Itza* have an additional visual component obtained by modifying the colour of the sky, which is customized for each IL.

Figure 6: Chichen Itza

These two environments represent interesting test-beds, as they are visually and acoustically different. While *Newton's Cenotaph* is a highly symmetric building where different areas can hardly be distinguished without additional information, *Chichen Itza* is more similar to a real town where different areas are characterized by a specific morphology. The two environments also exploit different types of audio. *Newton's Cenotaph* is enriched with fragments of modern music, *Chichen Itza* contains natural sounds, such as waterfall. In both environments, however, audio characterizes only the spatial locations mapped with IL; rapid fade in and fade out happen when the visitors enter or exit the IL.

Three groups of users participated in the evaluation of both worlds. The control group used them without any audio component. A group of users used the IL enriched version, without any explanation on the audio role in the interaction. A group of participants used the IL enriched version as the previous group, but they were provided a comprehensive explanation of the IL concept and the nature of the audio component in it.

4.1 Method

Participants. Participants were 21 students of a multimedia design course at the University of Venice, Italy. They ranged in age between 20 and 25 years old, had homogeneous basic skills in the use of personal computers, and no professional experience related to cultural heritage.

Procedure. Participants were tested in a group setting. Before the evaluations, they were briefly informed about the nature of the environment they were going to evaluate and randomly assigned to one of the three experimental conditions (Unimodal, Informed, and Uninformed). Then participants used the guided tour option (see Figure 7) to visit in a sequence a selection of ILs without any control over the navigation. They were invited to pay attention to as many details as possible. The experimental phase consisted of a set of scene recognition tasks. Participants were automatically transported to a number of ILs (40 for *Newton's Cenotaph* and 38 for *Chichen Itza*) and asked whether they had already been there during the guided tour (see Figure 8); three options were available, Yes, No, Uncertain. Both visited and unvisited ILs were presented to the users. All participants tested *Newton's Cenotaph* first and *Chichen Itza* second.

Figure 7: Chichen Itza – Guided Tour

Design. Evaluation (3) was manipulated between-subjects. Participants were randomly assigned to one of three evaluation settings: Unimodal (no audio provided); Uninformed (IL with no further explanation) and Informed (IL with explanation of the audio role).

Figure 8: Chichen Itza - Experimental Phase

4.1.1 Results

The evaluation concentrated on user's performance in scene recognition. The graph in Figure 9 illustrates the percentage of wrong answers, correct answers and uncertain answers as a function of visited and unvisited ILs in the three evaluation conditions. The graph summarises results from the *Chichen Itza* evaluation.
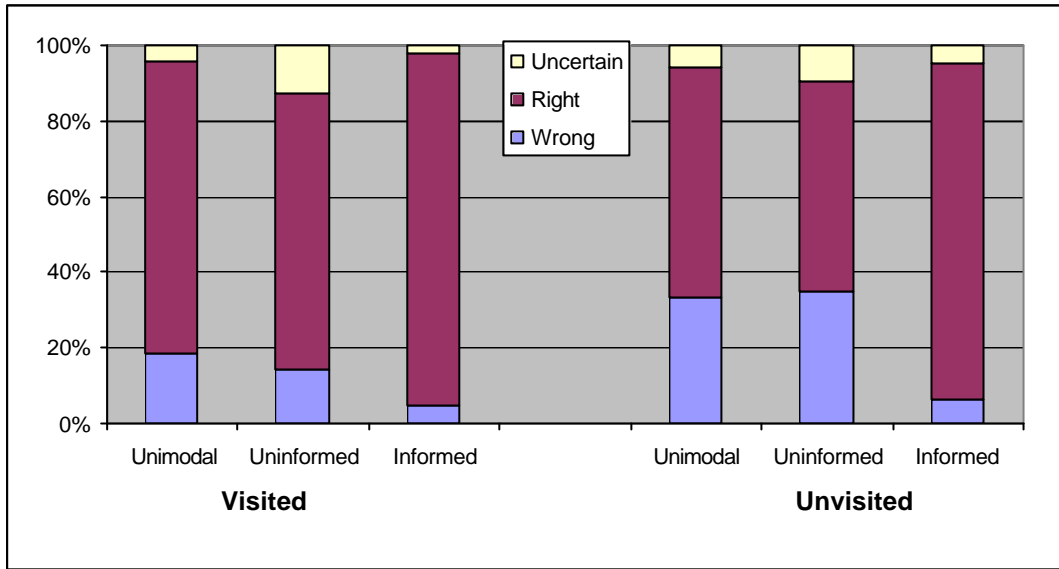
Figure 9: Scores of the scene recognition test in the Chichen Itza evaluation

The graph in Figure 10 illustrates the percentage of wrong answers, correct answers and uncertain answers as a function of visited and unvisited ILs and experimental conditions for the *Newton's Cenotaph* evaluation.
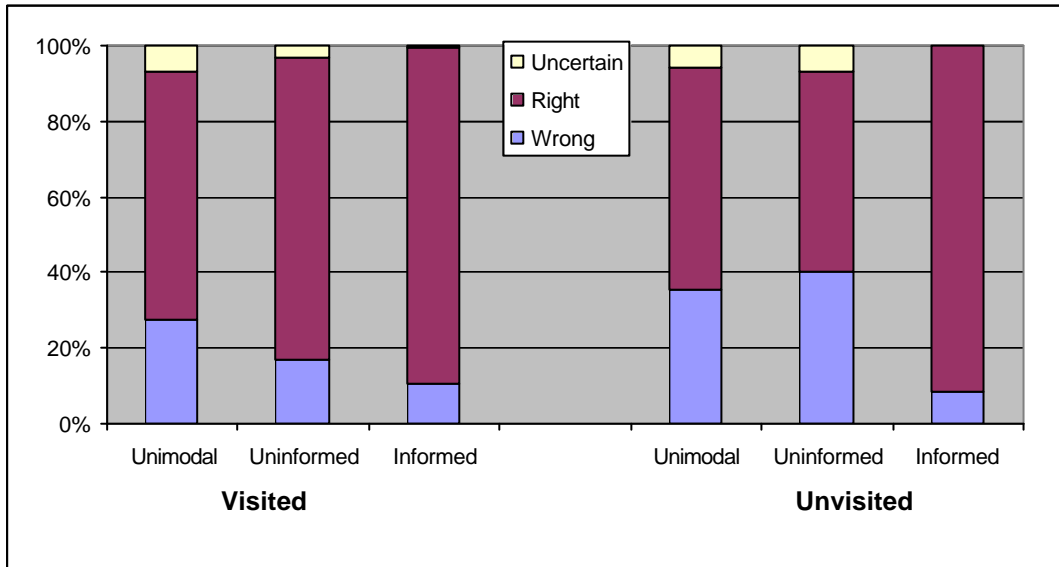

Figure 10: Answers to scene recognition test in the Newton's Cenotaph evaluation

The trend of results is very similar between the two applications. In the Informed conditions participants tended to be more often correct, than in all other conditions, independently of the 3D world evaluated. To test the strength of this effect an accuracy index was computed for each participant, by applying the following formula:

$$Accuracy = (Hits + Correct\ rejection) / Task\ Number$$

Hits are correct answers as regards visited places (the user has been there and say 'yes'), *correct rejection* regards unvisited place (the user has not been there and say 'no'). The accuracy index ranges from 0 to 100 (maximum accuracy).

The accuracy index was analysed by an ANOVA with Evaluation Type (3) as the between-subjects factor. Separated analysis were run on each evaluation data set. The effect of Evaluation Type was

significant in both cases, $F_{(2,20)} = 9.97$, $p < .01$ for *Newton's Cenotaph* and $F_{(2,20)} = 9.75$, $p < .01$ for *Chichen Itza*. Post Hoc tests suggested that the effect was mainly due to the Informed condition, which constantly achieved the highest accuracy in both applications. No significant differences emerged between the Unimodal and the Uninformed condition. People aware of the IL role performed significantly better than other participants did. People using IL enriched worlds without explanations of it did not perform any better than participants using the unimodal version.

4.1.2 Discussion

This study confirmed that mapping 3D environments with a set of **L**s provide significant navigational advantages, helping people to orient and recognize physical places. However, a plain implementation of this technique does not guarantee any benefit; the power emerges only when users are aware of the intended use of sound.

5. Conclusion

The goal of this work was to investigate the role of location-based hypermedia for helping navigation in 3D environments. Our main question in carrying out the experiments was whether the IL audio component could simplify navigation in VE. The answer is more subtle than expected. The first experiment showed that music could play an aesthetic role that users appreciate very much. This is an important result since user satisfaction is a significant component of system usability [13]. However, it did not provide any support to the IL assumption claiming a functional role of music in supporting navigation. Despite these not encouraging results, the experiment provided some interesting insights into the way in which people experience and respond to earcons in 3D worlds. It suggested a basic problem in understanding the exclusive link between sound and physical environment and in taking advantage of this one-to-one relationship to improve navigation.

To overtake this problem, in the two successive studies we used a simple experimental artefact, informing users about the music-space relationship. Results showed a significant improvement under this condition. When participants were informed in advance of the semantic link between audio and virtual space, audio helped navigation in virtual environments. Informed participants navigated faster and more efficiently, and recognised virtual scenes much more easily and accurately that if they were unaware of the link. Uninformed users were not capable of exploiting the auditory cues and performed no better than using a unimodal system, with no audio at all.

It is clear that audio is a complex medium and that the Interaction Locus is not a walk-up-and-use interaction technique, because users cannot autonomously understand the role played by earcons. Initially, we thought that this problem was due to a particular and very sophisticated type of classical music tested in the first evaluation, but the results remained consistent even when natural sounds where used as in the case of the *Chichen Itza* world. People tended not to associate sounds to location unless prompted. Further research is required to find better ways of exploiting IL techniques, building a sound semantic and understanding associations between sounds and objects as well as sounds and virtual locations. This effort appears worthwhile if one considers the intrinsic navigation advantages related to this approach.

References

[1]   M. Blattner, D. Sumikawa & R. Greenberg, Earcons and icons: their structure and common design principles. Human Computer Interaction, 4:11-44, 1989.

[2]   D. Bowman, E. Kruijff, J. La Viola & I. Poupyrev, The Art and Science of 3D Interaction. Tutorial notes from the IEEE International Virtual Reality 2000 conference. New Brunswick, NJ, March 18–22, 2000.

[3]   S. A. Brewster, P.C. Wright, & A.D.N. Edwards, An evaluation of Earcons for Use in Auditory Human-Computer Interfaces. In S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel, & T. White (Eds.), Proceedings of InterCHI'93, Amsterdam: ACM Press, Addison-Wesley, pp. 222-227, 1992.

[4]   S. A. Brewster, P.C. Wright, & A.D.N. Edwards, The design and evaluation of an auditory-enhanced scrollbar. In B. Adelson, S. Dumais, & J. Olson (Eds.), Proc. of CHI'94, Boston, Massachusetts: ACM Press, Addison-Wesley, pp. 173-179, 1994.

[5]   J. L. Chen & K. M. Stanney, A theoretical model of wayfinding in virtual environments: Proposed strategies for navigational aiding. Presence: Teleoperators and Virtual Environments, 8(6), 671-685, 1999.

[6]   F. Costalli, L. Marucci, G. Mori & F. Paternò, Design Criteria for Usable Web-Accesible Virtual Enviroments. D. Bearman, F. Garzotto, Eds., Proc. International Cultural Heritage Informatics Meeting (ICHIM) 2001, Cultural Heritage and Technologies in the Third Millennium, Milano, September 3-7, 2001.

[7]   A. Dix, J. Finlay, G. Abowd & R. Beale, Human-Computer Interaction. Hemel Hempstead, Hertfordshire: Prentice Hall Europe, 1998.

[8]   D. Fogli, F. Pittarello, A. Celentano & P. Mussio, Context-aware interaction in a mobile environment. Fifth International Symposium on Human Computer Interaction with Mobile Devices and Services (MobileHCI03), Udine, Italy, September 8-11, 2003.

[9] J. L. Gabbard & D. Hix, A taxonomy of usability characteristics in virtual environments. http://csgrad.cs.vt.edu/~jgabbard/ve/taxonomy/. 1997.

[10] W. Gaver, The SonicFinder: An interface that uses auditory icons. Human Computer Interaction, 4(1), pp 67-94, 1989.

[11] W. Gaver & R. Smith, Auditory icons in largescale collaborative environments. In D. Diaper et al. (Eds.) Human Computer Interaction - INTERACT '90, Elsevier Science Publishers B.V. (North Holland), pp 735-740, 1990.

[12] W. Gaver & R. Smith & T. O'Shea, Effective sounds in complex systems: the ARKola simulation. CHI'91 Conference proceedings, Human Factors in Computing Systems, "Reaching through technology" , New Orleans, pp 85-90, ACM Press: Addison-Wesley, 1991.

[13] ISO 9241: Ergonomics Requirements for Office Work with Visual Display Terminal (VDT), 1998.

[14] R.S. Kalawsky, The science of virtual reality and virtual environments. Wokingham, England: Addison-Wesley. (1993).

[15] K. Kaur, Designing virtual environments for usability. Unpublished doctoral dissertation. City University, London, 1999.

[16] Palazzo Grassi, Palazzo Grassi web site. http://www.palazzograssi.it. 1997.

[17] F. Pittarello & M. Pittarello, G.F. Italiano, Architecture and digital exhibitions - The Einstein Tower world, in M. Goebel, J. Landauer, U. Lang and M. Wapler, Eds. Virtual Environments '98 Springer Verlag, Wien, 1998.

[18] F. Pittarello, Desktop 3D interfaces for Internet users: efficiency and usability issues, PhD Thesis, 2001.

[19] F. Pittarello, Multi sensory guided tours for cultural heritage: the Palazzo Grassi experience. D. Bearman, F. Garzotto, Eds., Proc. International Cultural Heritage Informatics Meeting (ICHIM) 2001, Cultural Heritage and Technologies in the Third Millennium, Milano, September 3-7, 2001, Vol. 1 pp. 73-90, 2001b.

[20] F. Pittarello, Accessing Information Through Multimodal 3D Environments: Towards Universal Access. UAIS Journal, 2(2):1-16, 2003.

[21] K. M. Stanney & M. Zyda, Virtual Environments in the 21st Century. In Handbook of Virtual Environments, Stanney K. M. Ed, Lawrence Erlbaum Associates, Mahwah, N.J., 1-14, 2002.

[22] D. Sumikawa, Guidelines for the integration of audio cues into computer user interfaces, Lawrence Livermore National Laboratory Technical Report, UCRL 53656, 1985.

[23] D. Sumikawa, M. Blattner, K. Joy & R. Greenberg, Guidelines for the syntactic design of audio cues in computer interfaces, Lawrence Livermore National Laboratory Technical Report, UCRL 92925, 1986.