PROPOSAL FOR THE 2010 COLLABORATIVE RESEARCH INITIATIVES:

NIEVE:
Navigation and Interfaces in Emotional Virtual Environments

Table of Contents

1. NIEVE

Name, acronym and abstract of the proposal
NIEVE: Navigation and Interfaces in Emotional Virtual Environments

2. Coordinator

Name of person in charge (belonging to an INRIA project-team)

George Drettakis, EPI REVES, CRI Sophia-Antipolis Méditerranée

3. Participants

Identification and roles of the participants (skill areas, points in common and complementary skills)

Partner 1: REVES INRIA research group: http://www-sop.inria.fr/reves & DREAM service INRIA Sophia-Antipolis

The coordinator of the NIEVE project will be George Drettakis, leader of the REVES group in Sophia-Antipolis. REVES specializes in rendering for virtual environments with sound. The main focus of REVES is the development of plausible rendering solutions, including important visual effects such as shadows and texture, and rendering of complex auditory environments, including up to thousands of sound sources. REVES publishes regularly at all the prime conferences and journals in computer graphics (SIGGRAPH, Eurographics, I3D, EGSR etc.); recent projects include an interactive global illumination algorithm by localizing visibility computation, a fast frequency domain approach to contact sound synthesis and a novel model for procedural noise (used for texturing) based on Gabor kernels.

REVES is at the Sophia-Antipolis research unit of INRIA. REVES will bring the graphics rendering experience in realistic and expressive rendering, as well as extensive experience in 3D audio.

Key Persons for the Project

Dr. George Drettakis (PhD, Toronto 1994) has published many papers in graphics, in particular on global illumination, visibility, relighting and augmented reality and interactive rendering at international conferences and journals. He was project coordinator for the ESPRIT project SIMULGEN (Phase I) and locally responsible for Phase II, and actively part of the ARCADE project. He was the person in charge of the INRIA participation of the CREATE IST project. From 2005-2008 he was the coordinator of the FP6 FET OPEN project CROSSMOD. He has been part of program committees of many international conferences, has co-chaired the papers program of EGWR ’98 and the Eurographics’2002 and 2008 Conferences; he is papers chair for SIGGRAPH Asia. He is an associate editor for IEEE Transactions on Visualization and Computer Graphics. He won the Eurographics Outstanding Technical Achievement Award in 2007, and is a Eurographics Fellow.

DREAM service Sophia-Antipolis. In addition to the REVES group at Sophia-Antipolis, the research engineer service DREAM will participate by providing software and hardware support for the activities to be performed in the Immersive Space. This will include some software development (when this corresponds to development of common interest), and the partial supervision of the research engineer requested via the ADT (Sect. 10) and the engineering intern (see below).

Dr. Jean-Christophe Lombardo (PhD, Grenoble 1996) is a research engineer, in charge of the INRIA Sophia Antipolis-Méditerrannée Immersive Space. After doing his thesis in computer graphics, he focused on virtual reality. As a post doctoral fellow, he designed, together with the Belgian company Barco, an immersive display for the Elf/Sanofi group. This prototype became the Baron workbench. He was then project coordinator for the INRIA Joined Initiative AISIM on laparoscopic surgery simulation. He worked for 10 years at the Centre Scientifique et Technique du Bâtiment where he designed, built and coordinated the Le Corbusier reality centre. He was the person in charge of the CSTB participation in the CREATE IST project. He joined INRIA Sophia-Antipolis to design the new virtual reality facility.

Partner 2: IRCAM – Acoustic and Cognitive Spaces Team

The research and development activity carried out by the Acoustic and Cognitive Spaces Team of IRCAM concentrates on the analysis, reproduction and synthesis of 3D sound environments and is based on signal-processing, acoustics and psychoacoustics. Beyond the music domain, this research is applied in fields of audio post-production, communication and virtual reality. One axis of research concentrates around technologies, models and applications linked to the introduction of virtual or augmented reality environments in everyday life. In this framework, virtual reality is exploited to immerse a user in environments in which the interaction between different sensory modalities is controlled and manipulated. Its unique features and flexibility are used to study the perception-action loop and central processes of space-related multisensory integration. Applications to health care involving 3D sound are developed in collaboration with various hospitals.

Key Person for the Project

Dr. Isabelle Viaud-Delmon (PhD in 1999) is a CNRS research scientist in behavioural neuroscience, and will be responsible for the IRCAM contribution to the project. Her field of research is dedicated to the study of multisensory integration and spatial orientation both in neurological and psychiatric patients, as well as normal subjects, with virtual reality techniques. She was coordinator of the project "Multisensory information processing in virtual reality: application to psychopathology" supported by the French program "Cognition and information processing" from the CNRS, grant CTI 01-54, and a principal investigator for CNRS of the FP6 IST Open FET project CROSSMOD. She co-chaired the International Multisensory Research Forum 2002.

Partner 3: BUNRAKU INRIA Research group

The BUNRAKU team (formerly SIAMES) has been conducting research for more than 15 years in the field of Virtual Reality interfaces and interaction techniques. This research activity explores the multiple modalities of multi-sensory interaction with virtual worlds, through the study and use of several types of information, e.g., visual (immersive visual peripherals), haptic (tactile or force-feedback devices), and auditory (spatialized sound). BUNRAKU team in INRIA has thus developed international expertise in the field of virtual reality and interactions between real and virtual humans. BURNAKU is notably the co-inventor of the concept of "pseudo-haptic feedback" with a first publication at IEEE VR in 2000 [Lécuyer et al., 2000]. Since 2000, INRIA validated this concept through numerous experimental and perceptual studies. INRIA also promoted and developed applications of pseudo-haptics in various fields, such as for medical simulation, vocational training, or graphical user interfaces, etc. All these results were systematically published in numerous papers in reference conferences and journals (IEEE VR, ACM CHI, Presence, ACM TAP, Haptics symposium, Eurohaptics, etc).

Key Person for the Project

Anatole Lécuyer is a research scientist at the French National Research Institute for Computer Science and Control (INRIA), in Rennes, since 2002. He received his PhD in computer science from the University of Paris XI, 2001. His main research interests include virtual reality, 3D interaction, haptic interaction and brain-computer interfaces. He is an expert in Virtual Reality for European Commission (EC) or French National Research Agency. Lécuyer is or was member of International Program Committees of major international conferences (World Haptics, Eurohaptics, IEEE 3DUI, ACM VRST, IPT-EGVE, etc) and organizer of tutorials and workshops in major VR events (IEEE VR, Eurohaptics, etc). Lécuyer is associate editor of the ACM Transactions on Applied Perception. He is the principle investigator of several national and international grants, including the European Project Natural Interactive Walking and the French projects OpenViBE1 and OpenViBE2.

4. Goals

Our overall motivation is to further our *understanding* of navigation in virtual environments (VEs), with the ultimate goal of developing and evaluating improved interfaces for navigation in immersive environments. This is a very broad and ambitious research topic, and in the context of a collaborative research initiative, we can only reasonably hope to concentrate on a few specific sub-problems, which we describe next.

As we will see later, there is evidence of significant overlap in brain structures related to spatial memory and orientation and that of emotion. Thus the *first question* we will examine in NIEVE will be that of the *modulation of emotional valence*[1] in immersive audiovisual virtual environments. We are particularly interested in the influence of high-quality 3D audio on emotion modulation. We believe that the bimodal (visual and auditory) aspect of VE (Virtual Environments) has been overlooked in the past; in addition the link with emotional responses in VEs has rarely been studied to our knowledge (Viaud-Delmon et al 2006). Our study will focus on the phobia of dogs as a way to modulate emotion in audiovisual VEs.

From a more algorithmic perspective, navigation in VEs involves the use of different views, i.e., egocentric ("first person") and allocentric ("top view") views during navigation tasks. Thus *the second question* we will study in NIEVE is that of *appropriate visual representations* for each view (for example, abstract "map-like" rendering for top views, and realistic views for first-person views), and, more interestingly, *appropriate transitions and combinations* between the different views. The study of egocentric and allocentric views has received much attention in neuroscience research and in VR/HCI (Virtual Reality/Human Computer Interaction) research to a certain extent e.g., (Guiard, Beaudouin-Lafon et al. 1999). Nonetheless, we are unaware of the study of these transitions and their consistency in the context of the appropriate visual and auditory representations in VEs (e.g., realistic vs. expressive styles (maps, sketches etc.), realistic vs. exaggerated or stylized sounds).

We are interested in investigating the above questions in a fully immersive setting, which is now provided by the 4-wall iCube recently installed at INRIA Sophia-Antipolis Méditerranée. This will be the first true research project to use this ambitious new immersive platform, which is in many ways unique (see Annex 1).

We will next develop novel *immersive navigation interfaces* based on the study of representations and views discussed above. The study of emotion will serve as an important and innovative methodology to evaluate the efficacy of the interfaces developed. In particular, we will develop an appropriate methodology to evaluate navigation interfaces in stressful environments, based on the insight gained by the initial emotion modulation study in phobic settings. This novel methodology can be seen as a "stress-test" for navigation interfaces: if the navigation interfaces developed are successful even in stressful setups, they will definitely be successful under "normal conditions".

NIEVE is thus organized in three tasks. In a first phase, two tasks will be developed in parallel, i.e., study of emotional modulation and study of representations/views, both of which have relatively well-defined requirements. The third task, executed in a second phase, will build on the results of the studies by developing and evaluating novel navigation interfaces using emotion.

The overall goals of NIEVE can thus be summarized as follows:

- An in-depth study of the modulation of emotion in audiovisual VEs, focusing on the case of phobia of dogs in the iCube.
- The study of appropriate representations, view combinations/transitions and auditory feedback for effective navigation in VEs
- The development of novel immersive audiovisual interfaces for navigation in VEs

---

[1] Valence (psychology and neuroscience), referring to the emotional value associated with a stimulus; e.g., a familiar face can have positive valence.

- The development of an innovative emotion-based methodology for the evaluation of immersive navigation interfaces.

5. Research Activities

In this section we present detailed descriptions of the three tasks of NIEVE. Recall that Tasks 1 & 2 will start at the beginning of the project and will advance in parallel. Task three will start at the end of year 1 (see end of Sect. 6 for a detailed plan).

5.1. Task 1: Modulation of Emotional Valence in Audiovisual Virtual Environments

Introduction

Modulating emotional valence in immersive audiovisual VEs is a complex and difficult endeavour. In NIEVE, we will concentrate on the specific case of phobia of dogs (cynophobia). This case has two main advantages for our goals: first, phobia of dogs is related both to visual (aggressive look and behaviour of dogs) and auditory stimuli (growling and barking) and second, it is a common phobia, allowing deployment of the experimentation in a non-clinical setting. We will build on initial work developed previously by IRCAM and REVES, in the context of the EU project CROSSMOD[2]; the initial results involved the use of a semi-immersive setup (a "powerwall" at IRCAM), and were very promising.

In the context of NIEVE we will develop a much more in-depth and relevant study, since we will concentrate on the usage of a fully immersive "CAVE-like" iCube setup which has very recently become available at INRIA Sophia-Antipolis Méditerranée. The use of full immersion allows us to elicit much stronger emotional reactions, since the sense of presence and immersion in the VE is much higher.

Spatial disturbances are common complaints in pathological anxiety (Berthoz and Viaud-Delmon 1999) and there is a remarkable overlap between the brain structures involved in spatial orientation and spatial memory, and those involved in emotion. Spatial cognition has rarely been studied in anxious patients, but a few studies have proved so far that they demonstrate a deficit in way-finding abilities and topographical representation (Jones et al 1996; Kallai et al 1995). During navigation in a complex environment, anxious patients become lost more often and utilise fewer navigation points than normal control subjects. Moreover, map-like representations that encode metric distance relationships in allocentric coordinates (world-centered) are often found to be inaccurate in these patients. Therefore, testing navigation capacities of subjects sensitive to phobia is a good way to assess the effectiveness of our solutions to modulate the emotional valence of the environment; we will be using this idea to develop a novel way to evaluate interface design in Task 3.

Traditionally, virtual reality concentrates primarily on the presentation of high fidelity visual experience. We aim to demonstrate that adequately combining the visual and the auditory experiences provides a powerful tool to enhance sensory processing and attention modulation.

Our approach will allow the user to be progressively exposed to potentially frightening situations. In the proposed scenarios the stake would be to assess whether such progressive exposure could be based on the modulation of the relationship between visual and auditory cues.

In order to treat a phobia, the virtual exposure needs to evoke genuine reactions and emotions. In the case of cynophobia, a very important component of the anxiogenic stimulus is auditory: dogs barking and growling are very efficient in provoking emotional reactions in any individual. We also aim at

---

[2] http://www.crossmod.org

demonstrating that adequately combining the visual and the auditory experiences provides a powerful tool to modulate attention and to make the exposure more efficient.

This study is the first of it's kind in the level of realism, interactivity and immersion which is provided by the realism and fidelity of VE, for both audio and visual rendering. As a consequence, it will provide deep insight both in the understanding of the combined influence of audio-visual experience and in the consequences of emotional valence modulation for VE navigation tasks. Both goals are exciting new research directions; the former will advance state of the art in VR-based neuroscience and the latter will provide the basis for the development of the novel interface evaluation methodology outlined in Task 3. Both these aspects will provide multiple opportunities for a deep and rich collaboration between IRCAM and the two INRIA groups.

Background and Previous Work

Virtual Reality (VR) has been employed as an alternative for in vivo exposure for the treatment of different phobias for the past decade (see Riva, 2005; Parsons and Rizzo, 2008; Cobb and Sharkey, 2006). Specific phobias have been successfully treated with VR, including arachnophobia (e.g. Carlin et al, 1997). However, few of these procedures have been conducted with VR involving multiple sensory stimulations. Modulation of perceptual processing depends on sensory inputs from all modalities, and can potentially influence perception and behaviour in multiple ways. We have chosen to work on cynophobia because the acoustic aspect of this phobia is much more relevant than in some other phobias, and is therefore an ideal target to test the efficacy of auditory and visual environments to generate presence and emotion. Exposure-based treatments for cynophobia can use in vivo techniques, imaginal techniques, or hybrid techniques (Rentz et al, 2003). In order to treat a phobia, the virtual exposure needs to evoke genuine reactions and emotions.

Diverse audio-based applications have been implemented in the last few years, involving 3D sound. Because strong emotional reactions can easily be elicited through audition (Bremner et al, 1999), we want to fully exploit the potentiality of 3D audio to increase the realism and richness of the immersive environment. Humans are easily distracted by the locus of an irrelevant auditory event when their visual attention is focused elsewhere (Van der Lubbe and Postma, 2005). This implies that endogenous orienting does not suppress auditory exogenous effects, and that audition can therefore serve as a strong cue to induce emotional reactions in individuals immersed in a visual environment.

Figure 1: The 3 different dogs with different postures (lying, jumping and standing) within the 3 different environments (the city, the hangar and the garden).

Scenarios and methodology

During the different scenarios we propose in our virtual scenes, the participant will be gradually exposed to anxiogenic situations with dogs, using a variety of interactive animations to simulate the dynamic behaviour of the dogs relative to the environment and the participant himself. A typical scenario starts with a dog quietly walking far away from the subject, and ends with a more aggressive dog growling and barking when the participant approaches. We use three different textures to present our dog model in more or less frightening races of dogs: Malamut, Minpin and the commonly fearful Doberman, shown above.

We will use three virtual scenes (see Fig. 2 below). Each scene has a distinct ambient audio environment rendered through virtual ambisonic sources and binaural audio rendering. Using a number of predefined scenarios, the participant is asked to perform simple tasks, such as reaching a specific location in the scene or looking for a particular object. The experiment starts with a training session in a dog-free environment, so the participant can get used to the setup and learn to navigate easily. Afterwards, the participant experiences each scenario without any intervention of the experimenter to maximise immersion. The level of anxiety is measured by the experimenter at the beginning and at the end of a scenario, and a debriefing with the participant concentrates on the emotional reactions evoked by the key events of the scenes. The details of the scenarios are presented in Annex II.

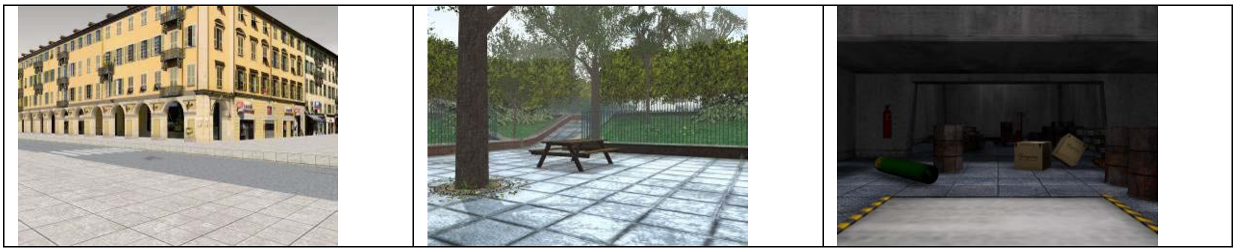


Figure 2: The three scenes: city, garden and hangar VEs

Evaluation sessions

Twenty participants selected for their high score on a dog phobia scale will be exposed to the various scenes following a strict procedure described below. The aim of this evaluation is to precisely assess the emotional reactions that are evoked during immersion. Each evaluation session will last about one hour, and will include an interview, a training immersion session without dogs, an immersion with dogs and a thorough debriefing.

Questionnaires and Interview Measures related to the exposure sessions

The state portion of the State-Trait Anxiety Inventory STAI (Spielberger et al 1983) will be used to measure the anxiety levels upon arrival at the laboratory and after completion of the exposure session. A 22-item cybersickness scale will be used to assess the level of discomfort immediately after exposure (Viaud-Delmon et al 2000). The “presence” questionnaire from the I-group (Schubert et al 2001) will be presented after immersion. Anxiety ratings will be collected during the exposure with Subjective Unit of Distress, a self-report measurement of anxiety on a 0–100 point scale (Wolpe 1973). The participant will be asked at beginning, at the middle and at the end of immersion about his/her level of anxiety. During exposure, the experimenter will rate the participant’s reaction in front of each dog.

Preliminary tests that were run in a semi-immersive initial prototype (CROSSMOD Deliverable D6.2 “Results of the Evaluations and Discussion of Applicability of Initial Findings in an Application Context”[3]), for the garden scene only, revealed very promising results. Indeed, our approach was shown to be very powerful in providing presence and immersion. First observations showed that the 3D sound component is promising in driving the progressive exposure of patients and their emotional reaction. In order to take advantage of the powerful role of auditory stimuli in the generation of fear reactions, the use of full VE immersion, including 3D audio (either with binaural rendering over headphones or a speaker based solution such as Crosstalk Cancellation) will be of central importance.

Several participants of our previous tests spontaneously reported that auditory stimulations were extremely arousing, and that they purposefully adapted their navigation to avoid walking with a growling or barking sound in their back. We expect such reaction to be even stronger in a fully

[3] This document can be provided if necessary, since it is not publicly available.

immersive setting of the iCube. The 3D sound component is therefore extremely promising to drive the progressive exposure of the participants, and to study navigation in a more ecological setting, providing multisensory involvement as well as emotional reactions.

Organization of Work for Task 1

Task 1 involves two main components:

1) experimentation performed by a psychologist and
2) development and content creation

The neuroscience experimentation will require the screening of a large number of candidates (we expect to screen around 100 candidates), to identify the set of 20 candidates to perform the experiment in the immersive environment. Once the 20 participants are selected, they will perform the 1 hour test, including the training and debriefing sessions. We will thus require a psychology intern, supervised by Dr. Viaud-Delmon; the work will be performed partly at INRIA and partly at IRCAM. All the experimental work will be done on site of INRIA. The preparatory work will be mainly done at IRCAM, in close collaboration with INRIA (preparation of the questionnaires to select the participants, analysis of the questionnaires, selection of participants, preparation of the interview and debriefing).

In terms of development, we will build on the initial application developed previously to fit the requirements of the three scenes presented above both in terms of audio and visual processing. The current application handles only the garden scene, with a limited number of interactive animations as a specific "hard coded" application. An appropriate scripting interface needs to be developed to allow the experimenter to be able to modify simple aspects of the interaction easily, and handle different scenes in a generic manner. This will require the development of a set of toolboxes for the design of such interactive environments, both for audio and visual processing. In addition, the current level of rendering quality is not sufficiently realistic for the experimental conditions: new algorithms need to be integrated into the VEs to handle dynamic lighting conditions and shadows, and better texturing.

These development tasks will be performed in part by an engineer who will be part of an ADT request in Sect. 9, and by an engineering student (M1 intern equivalent) to develop specific components required for the VEs of the tests.

In terms of content creation, both auditory and visual components will have to be designed. We will pay short-term task salary ("*vacations*") to a competent modeller with experience in virtual environments. These tasks include the development of more realistic animations, further dog models, improving and enhancing the existing environments, recording and manipulating sounds etc.

Both development and content creation will be co-supervised by J-C. Lombardo and G. Drettakis at INRIA Sophia-Antipolis.

5.2. Task 2: Representations and Views for Navigation

As mentioned previously, two kinds of representations are typically used in VEs for navigation: "map-like/top" views which correspond to the *allocentric* representation studied in neuroscience and "first-person" views which correspond to egocentric representations (McNaughton, Battaglia et al. 2006). In this task, we will investigate appropriate representations for each, inspired by the neuroscience literature and expertise of IRCAM in this domain. In addition, we will study the difficult question of smooth transitions between these views, which can only be achieved in the context of VEs.

There exist multiple forms of spatial knowledge: it is common to distinguish allocentric (examples are map- or top-views) and egocentric ("user-centered" or "first person") spatial information; also near (or peripersonal) from far (extrapersonal) space; and also front from back (non-locomotor) space. All these 'maps' have different neural correlates in the brain (Morgan 2003; McNaughton, Battaglia et al. 2006; Viaud-Delmon, Brugger et al. 2007). In real life, egocentric knowledge is viewpoint dependent, while allocentric knowledge is viewpoint independent.

In this task IRCAM will participate by providing knowledge of the neuroscience literature in spatial orientation, thus allowing the INRIA partners to make appropriate and efficient choices for representations and view combinations/transitions.

Study of Appropriate Representations

Despite the wealth of work on stylized rendering (see Fig. 3) (Strothotte and Schlechtweg 2002), there has been little research devoted to examining the relative effectiveness of different representations for different tasks. This is partly because such an evaluation is heavily application dependent. The context of NIEVE is a perfect opportunity to study this question in the specific context of navigation tasks.

Figure 3: Left: Realistic (direct lighting only) rendering and Right: pen-and-ink rendering (from [Coconu06]

There has been recent interest in "expressive" renderings for map-like views of cities in the context of the multiple virtual tourist applications such as GoogleMaps/GoogleEarth or VirtualEarth e.g., (Grabler et al 2008). Expressive rendering styles (Fig. 3), such as the Gooch NPR lighting model (Gooch et al 1998) or pen-and-ink styles (e.g., Coconu et al. 2006)) can be more appropriate for abstract representations of maps or streets when navigation is important.

We are particularly interested in determining criteria to choose a stylized rendering over a realistic rendering, and whether it makes sense to combine different styles. This will involve the development of novel algorithms for stylized, realistic and combined rendering. For realistic rendering, is it important to develop novel approaches, for example, for shadows or lighting, which adapt to these mixed representations. A fundamental question will be to investigate how to maintain both temporal and spatial coherence in the VE when some parts of the scene are rendered in the style of a map and some objects (e.g., important buildings) have a realistic illumination model, as well as how to achieve consistent transitions between the two. We will first investigate how to achieve this for prominent effects such as shadows and lighting, and then enlarge the study to other aspects of the image such as colour, material and texture or global illumination.

We will also investigate how to perform transitions between an expressive style and realistic styles, for example when moving from a map view to a first-person view, providing appropriate hybrid renderings at intermediate stages of the transition. We will use criteria e.g., such as facade saliency as outlined in (Grabler et al 2008) to determine appropriate transitions, as well as other measures of salient features in the geometry, texture and lighting used for realistic renderings.

Study of Views and Transitions

In addition to rendering style, we will investigate appropriate ways to handle the actual geometric/texture transitions involved in passing from a (typically 2D) map/top view to a (typically 3D) first person view. Evidently, the first step will be to use combined views (e.g., “map in a corner”) and abrupt transitions; we will then investigate the efficiency of different projections using them to smoothly transitions between views. An example could be oblique building projections (see figure below from (Grabler et al 2008) on a given map which would then “morph” into true 3D views in an expressive style, which subsequently blend into realistic rendering for the parts of the scene which are deemed important.

Figure 4: Oblique views (Grabler et al 2008), used in a “map view” which could “morph” into true perspective 3D when we move onto the street.

There is previous work on multi-perspective renderings and different projections style to achieve “detail-in-context” representations (e.g., Yu and McMillan 2004, Böttger et al 2008). These approaches distort the resulting image in different ways, depending on the goal to be achieved. The case of navigation will be inspired by these approaches, but we will have the additional concern of efficiency, since our solutions need to be real-time, and the selection of appropriate distortions in the goal of navigation. For example, it would intuitively make sense to distort a road on which we are traveling to achieve an exaggerated perspective view during transitions.

Auditory stimuli can also be manipulated to aid in navigation and in the transitions between views. The great advantage of 3D audio is that it can provide 3D spatial cues even with respect to parts of the scene (“non-locomoter space”) that become invisible once we leave the allocentric view. There is litot much literature, e.g., (Walter and Lindsay 2006) on audio cues to help in navigation; we will examine various modulations of the sounds used in VEs which will hopefully aid in navigation and in manipulating attention. Examples could include gain control, “non-realistic” manipulation of clustering and masking parameters (Tsingos et al 2004) to convey a sense of space or other spatial cues (for example a large crowd, even in a limited space) etc.

This task will also involve simple user tests to determine the effectiveness and efficiency of the various rendering styles and projection transitions. A set of specific tasks will be determined in given environments, allowing us to determine the effect of each individual choice as well as the combined effect. One goal of the user tests will also be to evaluate different conditions which will be required for the design of enhanced interfaces. This part of the work will be shared between Task 2 & Task 3.

Organization of Work for Task 2

The main part of the work for Task 2 will be performed by the postdoctoral fellow (see Sect. 8), who will spend 2/3 of her/his time on this task, and notably in the development of the novel algorithms and representations for view transitions and rendering. The postdoctoral fellow will also use the stable 3D audio library (APF) developed at REVES to perform the audio manipulations required for “non-realistic” audio effects.

The postdoctoral fellow will be based at REVES (Sophia-Antipolis), will visit BURNAKU and IRCAM for a few visits of 1-2 week and will supervised by George Drettakis. IRCAM and BUNRAKU will participate actively in the brainstorming sessions concerning this research. IRCAM will actively participate with their extensive knowledge of the neuroscience literature on spatial orientation, and BUNRAKU will contribute their experience in immersive VEs and navigation.

The design of the user studies to evaluate the projection and transition solutions will be the responsibility of the postdoctoral fellow in collaboration with BUNRAKU, however IRCAM and will be consulted regularly in the experiment design and analysis.

5.3. Task 3: Novel Interfaces using Emotion Modulation and Appropriate Representations

The third task will start at the end of the first year of NIEVE, and its goal is largely exploratory in terms of the study and development of interfaces based on initial results of the first two tasks. As such, it is necessarily shorter than Tasks 1 & 2 which will run for the full duration of the project.

Introduction

Current navigation interfaces for games and virtual environments typically involve either separate allocentric and egocentric views (an example is GoogleEarth, with “map view” and “Streetview”), or have a superposed map, often using some type of transparency, which appears in a corner of the first-person view. This approach works relatively well for non-immersive settings, but can be inappropriate in a fully immersive setting. The “world-in-miniature” concept has been proposed (Stoakly et al 1995) as a replacement to such a “superposed allocentric view” in immersive environments. However, it is unclear how successful such interfaces are for *immersive audiovisual* navigation tasks – this will be an interesting avenue of investigation.

Research Tasks

We will pursue two axes in Task 3 of NIEVE:

1. Novel audio visual interfaces using immersive allocentric representations
2. Emotion modulation to study navigation interface efficiency

The first axis will concentrate on different kinds of allocentric representations to aid navigation in immersive environments, in particular using different projections and transitions, rendering styles and audio attractors as developed in Task 2. The interface designs will include “planar maps” or WIM-like interfaces which are superposed in the environment, potentially using proprioception for access to the “widgets” (Mine et al 1997), or “pseudo-transparency” which can appropriately map an allocentric view to an egocentric context, based on appropriate projections developed in Task 2. We will design interfaces that appropriately combine rendering styles and transitions, and also investigate the use of sound as an attractor for navigation interfaces. One avenue of research is to study whether audio “beacons” (Walter and Lindsay 2006) are perceived as being less intrusive than visual signs, such as the “footprints” used for example in (Darken and Sibert 1993). Such “audio navigation widgets” will be an active part of the navigation interfaces developed.

The second axis involves the use of emotional valence modulation to efficiently test navigation interface designs. Based on the interface designs developed above, and the results on emotion modulation in Task 1, we will design specific user tests where the stress and fear levels of participants are increased progressively. The basic tenet of this approach is to “stress-test” interface design. The findings on Task 1 should reinforce our idea that emotional valence of the VEs deteriorates performance in navigation tasks; based on this (tentative) conclusion, we will see which parts of the interface design break down first when we increase the stress or fear levels of participants. Given the correspondences between the brain functions of navigation and emotion, we expect this methodology to provide very interesting insight into the design of immersive navigation interfaces, leading to better design in the long term.

We will develop navigation scenarios for one or two applications. Potentially scenarios include map-based navigation for adventure/first person like games, or a "guided tour" application for the presentation of a virtual city (for example a virtual visit of Cannes by a novice user).

Organization of Work for Task 3

The design and development of the immersive interfaces will be conducted by a Master's student co-supervised by REVES and BUNRAKU in close collaboration with the postdoctoral fellow will also participate in this Task with 1/3 of their time. The postdoctoral fellow will conduct the emotion modulation user tests in close collaboration with IRCAM, with support from the Master's student. The research engineer requested as part of the ADT will participate extensively in the design of the novel interfaces and also in theimplementation of the various audio-visual widgets (see Sect. 9).

5.4. Mutual Benefits of the Collaboration

NIEVE is centred on navigation in immersive virtual environments, and in particular the dominant role of emotion with a methodology founded in neuroscience results and competences. The mutual benefits of this collaboration are based on the fact that a multi-disciplinary team, interested in the same overall topic will work together to produce common results. Our experience with multidisciplinary research indicates that it is important to both maintain individual vested interest: in this case the study for phobia in VEs for IRCAM, rendering for REVES and interfaces for BUNRAKU while simultaneously including central collaborative aspects throughout. This is achieved via two main components: 1) the study of emotion and its central connection to spatial reasoning which affects all study of navigation 2) the expertise in spatial orientation in neuroscience which is brought to INRIA from IRCAM. These components permeate all three Tasks, and will guarantee high levels of collaboration.

The clear benefit for IRCAM in the context of neuroscience will be the ability to work with computer graphics and virtual reality experts in the context of a "CAVE-like" setup and thus a truly immersive environment. In addition, the supervision and monitoring of the VE development by the engineering team (J-C Lombardo, the ADT engineer and the interns) and the graphics/audio (REVES) and VR/HCI (BUNRAKU) researchers will provide an ideal capacity to pursue experiments with unprecedented levels of realism and interactivity (Task 1). The participation of IRCAM in the brainstorming, design and user-testing phases of Tasks 2 & 3, will be an interesting experience to researchers with a neuroscience focus; the use of emotion modulation for interface design opens an interesting direction for their more "core science" studies.

The graphics and audio partner (REVES), will benefit from this participation and guidance in the development of the various perceptually-motivated tools. The introduction of emotion modulation in the design of interactive algorithms will be a ground-breaking and novel experience for both INRIA groups (BUNRAKU, from a more HCI perspective, but also REVES). BUNRAKU will also benefit from the experience in audio and interactive rendering of REVES and the interaction with IRCAM. The HCI/VR experience of BUNRAKU will also be extremely beneficial to the other partners, in particular in the design of interaction metaphors and "widgets" (both for visuals and audio

6. Expected results

The major result of NIEVE will be to provide significant advances in the state of the art, first in the domain of the modulation of emotional valence in a neuroscience/VR context, second in the domain of rendering representations and transitions in the context of navigation and, third in a novel methodology of emotion-drive evaluation of novel immersive navigation interfaces. We expect to produce high-level publications in the three domains concerned, namely neuroscience/emotion, computer graphics and VR/HCI. We will target publications in the high level international conferences and journals for each domain. At the end of the two years of the project, we expect to have submitted or published at least one publication in each of these domains.

We are hopeful that the emotion modulation approach for evaluating navigation interfaces will have wider applicability, and potentially provide a more general methodology for "stress-testing" interactive 3D interface design. We believe that the potential impact for this approach is very significant, although we need to develop the research described here before being able to have a more precise understanding of the potential.

Another result will be the development of the demonstrators used for the Tasks. The demonstrator of Task 1 could provide the basis for the development of a therapeutic application for the treatment of phobias in a clinical setting. It will definitely provide an impressive showcase for the research developed here, and for the Immersive Space at Sophia-Antipolis. We also hope that the prototype interface designs (Task 3) will be potentially useful in the context of immersive environments.

7. Postdoctoral Researcher

The goal of this postdoctoral fellowship is to study appropriate rendering representations for navigation, and to use them to develop appropriate audiovisual navigation interfaces in immersive environments. The fellowship has two parts, first concentrating on representations and views for navigation (12 months) and a second part concentrating on immersive navigation interfaces (6 months)

The first part of the fellowship will concentrate on the study of realistic vs. non-photorealistic rendering styles for navigation tasks and on appropriate combinations or transitions between "top" or "map view" and "first person" views.

We are particularly interested in determining criteria to determine when a stylized rendering should be chosen over a realistic rendering, and whether it makes sense to combine different styles. The study of such questions will involve the development of novel algorithms for stylized, realistic and combined rendering. A fundamental question will be to investigate how to maintain both temporal and spatial coherence. In addition to rendering style, we will investigate ways to combine, and transition between, top- or map-views which typically use expressive rendering styles, to first person views which involve realistic or a combination of realistic and expressive styles. 3D audio may potentially be used to facilitate these transitions.

The second part of the fellowship will build on the appropriate representations developed previously to build novel navigation interfaces in an immersive environment (the Immersive Space and INRIA Sophia-Antipolis). The interfaces will be evaluated in several scenarios, and will also be part of a study of the influence of emotion on navigation in Virtual Environments.

The postdoctoral fellow will be assigned to REVES, but will spend time both at IRCAM and BUNRAKU.

8. Technological Development

We will be submitting an ADT in 2010, jointly with the DREAM service. The position will have a 50% assignment to the NIEVE ARC and 50% to the DREAM service for the development of software tools of general utility for the Sophia-Antipolis research center.

In what concerns NIEVE, the software engineer will collaborate with the engineering intern, to extend and adapt the initial application developed previously to fit the requirements of the three scenes presented above both in terms of audio and visual processing, and to fit the immersive space (Task 1). An appropriate scripting interface needs to be developed to allow the experimenter to be able to modify simple aspects of the interaction easily, and handle different scenes in a generic manner. This will require the development of a set of toolboxes for the design of such interactive environments, both for audio and visual processing. In addition, the current level of rendering quality is not sufficiently realistic for the experimental conditions: new algorithms need to be integrated into the VEs to handle dynamic lighting conditions and shadows, and better texturing. In addition, the software engineer will work with the Masters intern and the postdoc in Task 3 for the development of novel interfaces for navigation, using the results of realistic vs. expressive rendering and the view transitions/combinations developed in Task 2.

9. Appendix 1 – INRIA Sophia-Antipolis Immersive Space

The Immersive Space in Sophia-Antipolis has two display devices, a CadWall for its ease of use and an iCube for its immersion quality. Both devices offer 6:1 sound restitution, stereo images and optical tracking. In NIEVE, we will be using the iCube.

iCube

This display device offers a 3.2m x 2.4m working area. It is made of 3 vertical screens (left, front, right) and an horizontal one (floor). Thanks to the head tracking, it allows a single user to be completely immersed in a virtual environment.
It is a unique setup at INRIA since it is the first immersive “CAVE-like” immersive environment of the Institute; in particular, the floor is retroprojected, which is essential for the sensation of immersion and presence. In addition, we provide full Head-Related Transfer Funtion based 3D audio spatialization, delivered over headphones or speakers using Cross Talk Cancellation (planned for 2010).

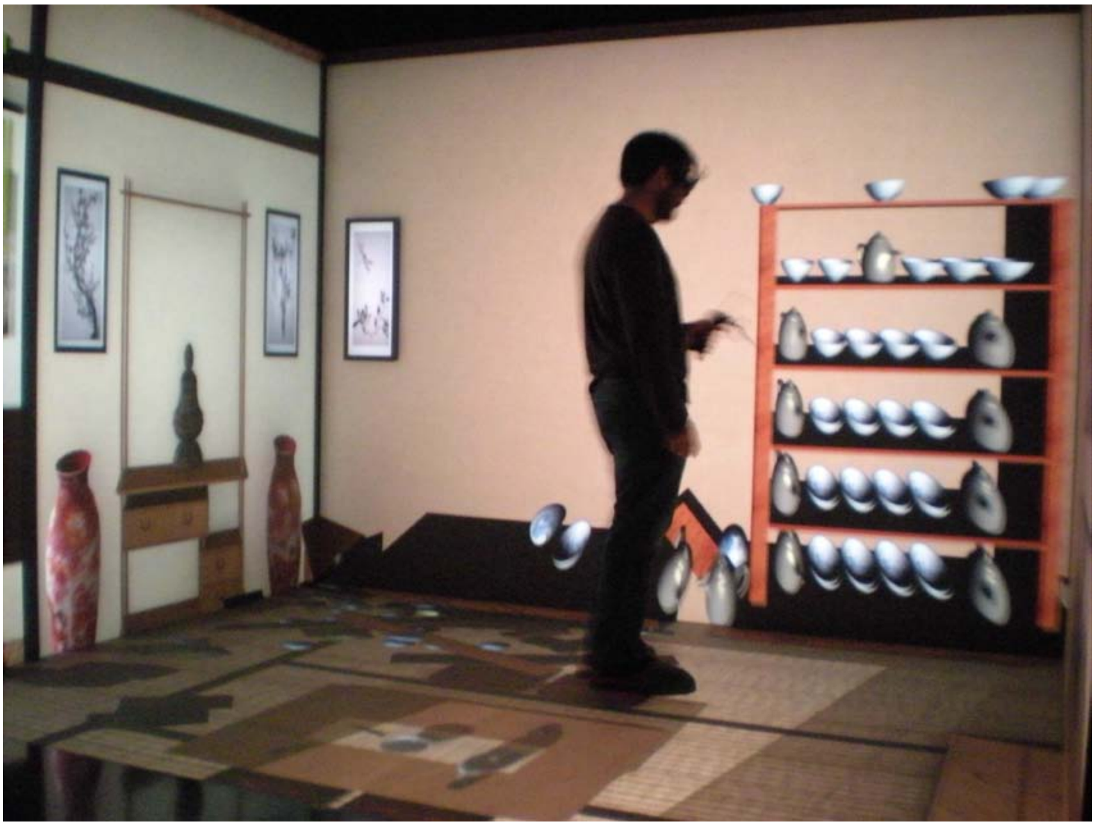
Technical features

Enforcing the user experience quality has been one of our main goals. It results in several features:

- Screens
  - Retro projected screens (including the iCube floor) to avoid shadows.
  - Dark screens to enhance contrast.
- Projectors
  - 12k lumen to guarantee brightness.
  - Xenon bulbs for their spectrum.
- Stereo
  - Active Infitec to minimize ghosting and avoid synchronization issues

Two different image and sound generators are available:
- a single PC with multiple graphics card and a multi canal sound card,
- a graphics cluster where a master drives one slave dedicated to each projector and a sound server. Graphics cards are Nvidia Quadro 5800FX.

A photograph of the iCube is shown below:



10. Appendix II: Scenarios for Dog Phobia

The first scenario takes place in a city. It is a wide open scene in daylight, composed of a large square surrounded by buildings. The ambient audio environment is a familiar city sound with distant cars and rumours. Among the buildings, there are a number of locations of interest such as shops, banks or restaurants. The scenario consists of a walk around the city, where the participant is asked to successively reach three different places: a restaurant, a supermarket and a cash dispenser. On his way through the city, he will encounter 4 dogs.

The garden scene is a rich, open environment, with dense vegetation surrounding two large houses. The scene takes place in daylight, with a light fog floating around, increasing the feeling of mystery of the garden. The ambient audio environment is a peaceful garden sound with singing birds and distant voices. The objective of the participant in this environment is to find a frog hidden in the garden, which is both a visual and auditory object. Progressive exposure to dogs is achieved by walking through the environment from an initial square to a final square, via a narrow passage way. Dogs appear first behind bars and later in front of the subject.

The hangar scene is an interior, dark scene, composed of a main rectangular room with entrances to smaller rooms at both sides. It features a rich auditory environment, provided by the activity of industrial machines along the sides of the main room. The task of the participant is to walk across the main room in the direction of an automated door, walk through the automated door, which is regularly opening and closing. Once in the other room, he has to count the barrels and to walk back to his starting point. Again, the subject encounters dogs in varying conditions, allowing both auditory and auditory-visual exposures.

Bibliography

Berthoz A and Viaud-Delmon I (1999). Multisensory integration for spatial orientation, Current Opinion in Neurobiology, 9, 708–712.

Bremner J D, Staib L H, Kaloupek D, Southwick S M, Soufer R, Charney D S (1999), Neural correlates of exposure to traumatic pictures and sound in Vietnam combat veterans with and without posttraumatic stress disorder: a positron emission tomography study, Biological Psychiatry, 45, 806-816.

Böttger, J., Preiser, M., Balzer, M., Deussen, O. (2008), Detail-in-context visualization for satellite imagery, Proceedings Eurographics, Eurographics Association, 2008.

Brooks T, Camurri A, Canagarajah N, Hasselblad S (2002). Interaction with shapes and sounds as a therapy for special needs and rehabilitation, Proc. 4th Intl. Conf. Disability, Virtual Reality & Assoc. Tech., Veszprem, Hungary, pp. 205-212.

Carlin A S, Hoffman H G, Weghorst S(1997), Virtual reality and tactile augmentation in the treatment of spider phobia: a case report, Behavior Research and Therapy, 35, 153-158.

Coconu L., Deussen O., Hege H-C. (2006) Real-time pen-and-ink illustration of landscapes, NPAR 2006.

Darken R, Sibert J. (1993) A Toolset for Navigation in Virtual Environments. ACM Symposium on User Interface Software and Technology, Atlanta, Ga..

Gooch A., Gooch B., Shirley P. and Cohen E. (1998), A Non-Photorealistic Lighting Model For Automatic Technical Illustration (SIGGRAPH 1998).

Grabler F., Agrawala M., Sumner R. W., Pauly M., (2008) Automatic Generation of Tourist Maps, SIGGRAPH 2008

Guiard, Y., M. Beaudouin-Lafon, et al. (1999). Navigation as Multiscale Pointing: Extending Fitts' Model to Very High Precision Tasks. Proc. ACM Human Factors in Computing Systems, CHI'99, Pittsburgh (USA), ACM Press %U http://doi.acm.org/10.1145/302979.303128: 450--457.

Cobb SV G and Sharkey PM (2006). A decade of research and development in disability, virtual reality and associated technologies: promise or practice? Proc. 6th Intl. Conf. Disability, Virtual Reality & Assoc. Tech., Esbjerg, Denmark, pp. 316.

Jones RB, Humphris G, Lewis T. Do agoraphobics interpret the environment in large shops and supermarkets differently. Br J Clin Psychiatry 1996;35:635–7.

Kallai J, Koczan G, Szabo I, Molnar P, Varga J. An experimental study to operationally define and measure spatial orientation in panic agoraphobic subjects, generalized anxiety and healthy control groups. Behav Cogn Psychotherapy 1995;23:145–52.

Lahav O and Mioduser D (2002); Multisensory virtual environment for supporting blind persons'acquisition of spatial cognitive mapping, orientation, and mobility skills, Proc. 4th Intl. Conf. Disability, Virtual Reality & Assoc. Tech., Veszprem, Hungary, pp. 213-220.

Mark M., Brooks Jr. F, and Sequin C. (1997). Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction. Proceedings of SIGGRAPH 97, Los Angeles, CA.

Parsons TD and Rizzo AA (2008), Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis, Journal of Behavior Therapy and Experimental Pyschiatry, 39, 250-261.

Rentz TO, Powers MB, Smits JA J, Cougle JR, Telch MJ (2003) Active-imagianal exposure: examination of a new behavioural treatment for cynophobia (dog phobia), Behaviour Reasearch and Therapy, 41, 1337-1353.

Riva G. (2005), Virtual Reality in Psychotherapy: Review, Cyberpsychology & Behavior, 8, 220-230.

Sanchez JH (2004), AudioBattleShip: blind learners cognition through sound, Proc. 4th Intl. Conf. Disability, Virtual Reality & Assoc. Tech., Oxford, UK, pp. 199206.

Schubert T, Friedmann F, Regenbrecht H (2001). The experience of presence: Factor analytic insights, Presence Teleoper Virtual Environ, 10, 266−281.

Spielberger CD, Gorsuch RL, Lushene R, Vagg, PR, Jacobs GA (1983). Manual for the State-Trait Anxiety Inventory (STAI), Form Y. Palo Alto, CA: Consulting Psychologists Press.

Stoakley, R., M. Conway, J., et al. (1995). Virtual reality on a WIM: interactive worlds in miniature. CHI '95: Proceedings of the SIGCHI conference on Human factors in computing systems.

Strothotte, T. and S. Schlechtweg (2002). Non-Photorealistic Computer Graphics: Modeling, Rendering and Animation, Morgan Kaufmann

Tsingos, N., Gallo E., Drettakis G. (2004). Perceptual audio rendering of complex virtual environments. ACM SIGGRAPH 2004 Papers. Los Angeles, California, ACM.

Van der Lubbe RHJ and Postma A (2005), Interruption from irrelevant auditory and visual onsets even when attention is in a focused state, Exp Brain Res, 164, 464-471.

Viaud-Delmon I, Ivanenko YP, Berthoz A, Jouvent R (2000). Adaptation as a Sensorial Profile in Trait Anxiety: a Study with Virtual Reality. Journal of Anxiety Disorders, 14, 583-601.

Viaud-Delmon I, Warusfel O, Seguelas A. Rio E. Jouvent R. (2006) High sensitivity to multisensory conflicts in agoraphobia exhibited by virtual reality, European Psychiatry, 21, 501-508.

Walker, B., N. and J. Lindsay (2006). Navigation Performance With a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice. Human Factors: The Journal of the Human Factors and Ergonomics Society

Wolpe J (1973). The practice of behavior therapy (2nd ed.), Pergamon, New York.

Yu J., McMillan L. (2004), A framework for multiperspective rendering, Eurographics Symposium on Rendering, 2004