Multiple Presence through Auditory Bots in Virtual Environments

Martin Kaltenbrunner
FH Hagenberg
Hauptstrasse 117
A-4232 Hagenberg
Austria
modin@yuri.at

Avon Huxor (Corresponding author)
Centre for Electronic Arts
Middlesex University
Cat Hill, Barnet
UK
a.huxor@mdx.ac.uk

0. Summary

This paper proposes that virtual environments that aim to support mutual presence for distributed work groups should allow for multiple partial presence. The introduction of teleworking, and the massive uptake of mobile phones have addressed a need for people to be virtually present (if not physically present) from many spatial locations. Conventional virtual environment systems, however, seem fixed to a notion of presence as being tied to 'being in' a specific place. It is this view of VR that we have modified with collaborative working in mind.

Specifically, we have extended a graphical Internet-based 3D world to allow for users to have multiple, 'proxy', avatars which provides a sense of partial presence in many locations in the virtual world, and/or in many different virtual worlds. Presence is now not tied to place, but to awareness of events. Users are connected to their proxies by audio cues, allowing for multiple locations to be attended to at once, creating a form of presence appropriate to the workplace.

1. Introduction

In previous work (Huxor, 1999), an application of shared virtual spaces to support distributed working was described. It is known that 'chance encounters', the unscheduled meetings between people that occur in such places as corridors, are crucial to the functioning of organisations (Backhouse & Drew, 1992). It is these encounters that any collaboration tool must support in addition to the more formal aspects of collaboration. It was suggested that such encounters can be better implemented in a shared 3D space, as opposed to than other collaborative web-based environments due to the importance of spatiality in managing interpersonal encounter. This idea has been explored through an example prototype, which employs the ActiveWorlds (http://www.activeworlds.com) shared 3D-world technology, hyperlinked to work-related content held in BSCW (http://bscw.gmd.de), a web-based document and collaboration management system. Certain physical spaces, known as 'locales' contain content and other people that are associated with a particular work-task: an association which manages encounters, and equally virtual 'locales' can facilitate online encounter between distributed work-groups.

One author has used this system for over three years to support such distributed working practices. It has indeed facilitated chance encounters not only for myself, but also for other colleagues. Although the system proved valuable, it was found that many potential encounters with other users were missed, as either:

- I had filled my monitor screen area with another application that required the space, making the virtual world not viewable, or
- My gaze was directed elsewhere: at a paper document, at a visitor to my physical office, out the window, etc.

Such concerns, however, also point to a larger issue. Unlike many virtual activities (such as gaming or training simulations) collaborative working requires that onehas a variety of presences - ofawarenesses. One can be both in a physical office and engaged in a telephone conversation: in this situation where is one's presence? It seems to be in both 'places' at once.

The problems of missed encounters were addressed by the addition of audio cues, so that activity in the virtual world generates various appropriate sounds that can be heard when the user is unable to view the 3D world. A user can set up avatar representations in a number of locations in the virtual world, each one specific to an ongoing work task of interest. While otherwise engaged, this user can now follow much of the 'coming and goings' and other activity that is taking place in all of these locations. The identity of which location the action occurs being indicated by variations in the sound cue.

One of the surprising results is that moving away from a purely visual presence to an auditory one creates new possibilities of presence. Specifically, we have extended the system to allow for multiple 'proxy' avatars in various spaces. This would be impracticable with a conventional visual interface - one window only on the virtual world takes up enough screen real estate, multiple windows would make other work impossible.

It seems strange that as technology in the workplace tries to liberate us from being fixed to a single location, through the use of mobile for, for example, that collaborative virtual environments should try to re-impose that constraint. The use of audio, and its ability to carry multiple steams, allows us to exploit the potential of VR technologies and support partial presence in many locations at the same time. With this in mind, we have extended the audio features of AlphaWorld. A user can now work in the physical office, and need not even have the 3D-world window on-screen. They can set a number of probes, or proxies, of their avatar in a variety of places. Each place in the environment is associated with particular content, in that the various object in the location are hyperlinked to online content which is stored in BSCW. Most work consists of managing a number of tasks at the same time, dealing with a variety of projects. Just consider a typical day: writing a document, while reading and responding to new emails, phoning a colleague about an urgent problem, and so on. The AlphaWorld virtual office that we use follows this idea, creating task-related rooms and buildings in a larger social community.

Each proxy is a bot that generates suitable audio cues for important activities at each location, such as another user entering or leaving, another user text-chatting, or a hyperlinked object being activated. This audio bot system is called Marvin. Figure 1 shows a screen shot from the AlphaWorld office, one of the proxy bots is in the space, represented by the avatar of a bird (this avatar was selected from the limited number available in AlphaWorld due to its connotation of being somewhat 'above it all', of partial presence).

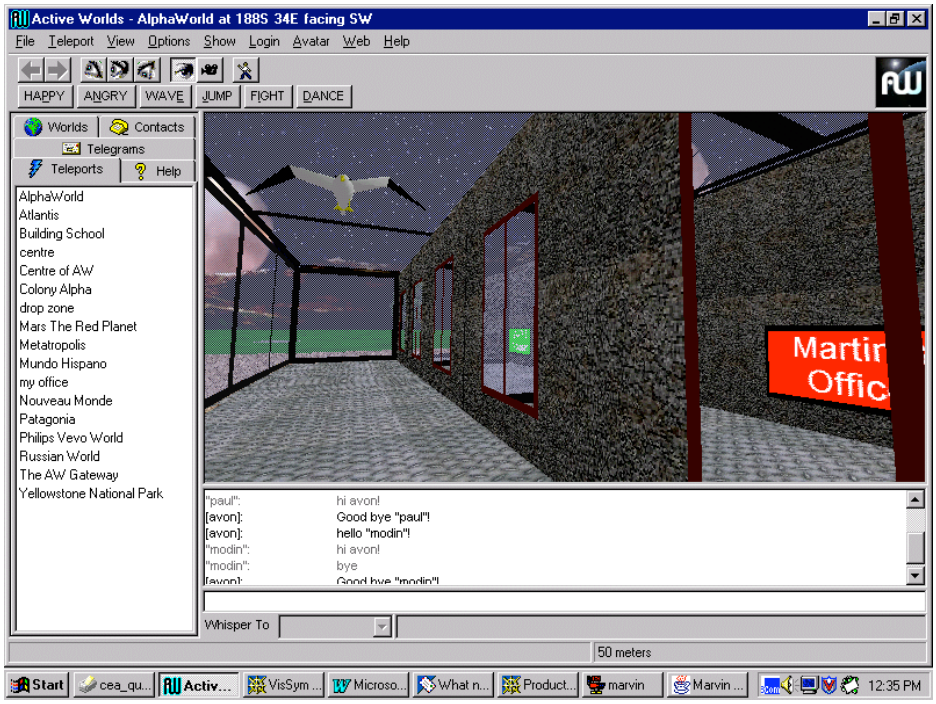


Figure 1. A Marvin proxy in the Virtual Office

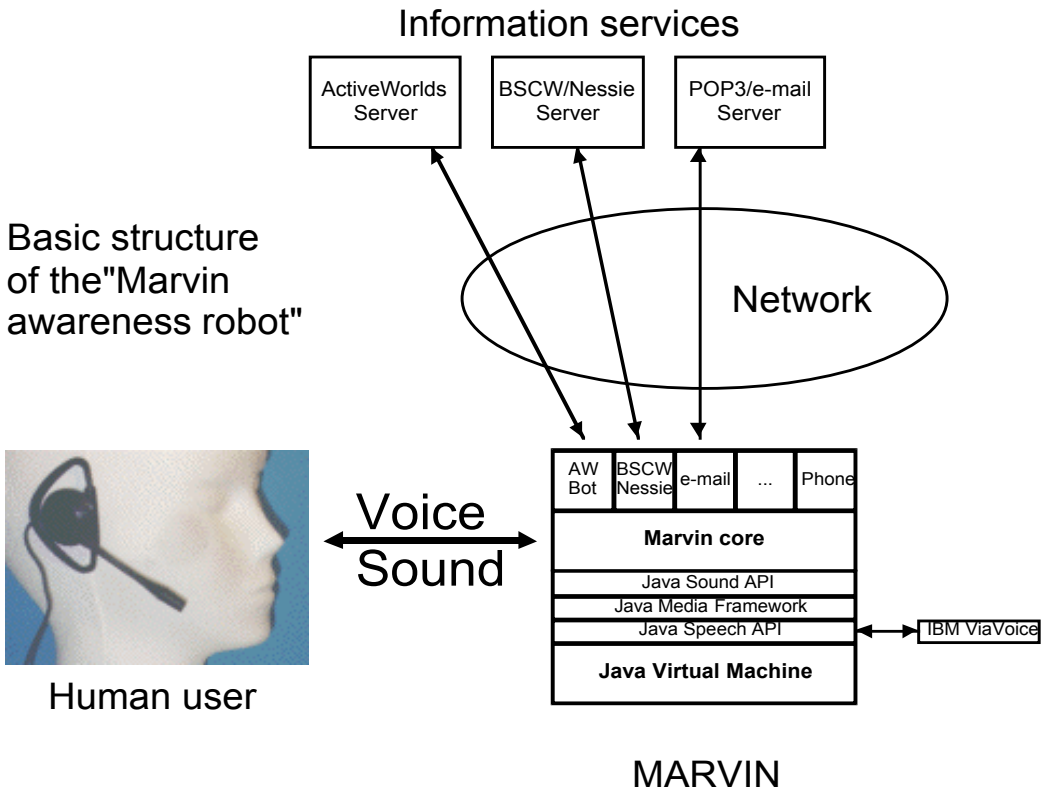
2. Marvin: an auditory bot

The Marvin bot is a simple programmable agent, who can enter multiple information servers (such as AlphaWorld) as a proxy and report noteworthy events via speech and audio to the user. Since with the help of the robot one is constantly aware what is going on, he can if desired then directly turn his attention to the appropriate application and enter the locale instead of the robot. The proxy itself isn't an independent intelligent agent, it is an extension to the user's senses for extended awareness of events in info-space.

Marvin is implemented in the Java programming language, version 1.1 or above. This decision was made mainly because of Java's platform independence and the availability of various multimedia features like Speech, Sound, and Media APIs. The application consists of two layers:

1. The Marvin core classes, which provide the basic features like speech recognition and synthesis, sound output, logging and so on. This core loads all present plug-ins after initialisation and starts them as independent threads. The plug-ins then can use Marvin's event processing methods. Depending on the event's priority, it will either only logged to a file or to the screen, or in case of the highest priority announced with speech and audio cues.

2. The plug-in interface, which allows the easy addition of any information service. At the moment we have only implemented an Active Worlds Robot. But due to this architecture and the fact that Marvin is released under the GPL (Gnu General Public License) it is easy for third party developers to add their own plug-ins for various other network information services or multi-user environments.



2.1 The Active Worlds Robot

The AWBot plug-in uses ThierryNabeth's Java native interface for the Active Worlds Software Development Kit. Applications implementing this interface can place a remote-controlled avatar into the Active Worlds space, which can interact with other persons or events in this virtual environment. Once such a robot is placed in a certain predefined area of a world within the AW server, it will notice any event in its surroundings. These events are then caught be the robot application, analysed and then processed by the Marvin core according to their priority. Since, as mentioned above, a robot only can notice events in its nearest surroundings, multiple instances of the robot – called probes – were created and placed in various areas of the AW space. To ease the modification of the robots, all the variable parameters such as position, user names or sound files are stored in separate configuration files that are processed upon start-up.

3. Discussion

We have used readily available Internet 3D world technologies to explore the use of desktop VR for collaborative working applications for some years now. Throughout that time one issue constantly arises. The technology has been primarily developed for certain simulation tasks, for training and gaming. The notions of immersion and presence which are appropriate for such applications differs greatly from the needs of CSCW (computer-supported collaborative working) in which the virtual environment is part of ongoing work activities. Indeed, it may be better to consider such systems as a form of augmented reality, as users always require 'one foot in the physical world'.

Experience of the use of the audio bot system has confirmed many of our intuitions: Not only of the importance of multiple presence, but of the power of the audio channel. The

response to sound is so very different to that of visual information. A virtual doorbell sounds like a physical one – both has the same form of waves in air hitting the ear. A virtual visual door image, however, does not yet have that same equality of experience. In future work we hope to explore what the nature of this audio mediated presence. Two possible factors seem worth investigating. Firstly, the role of social presence (as discussed by Towell & Towell 1997), that is, to what extent is a sense of being there sustained by the possibility of interaction with other (interesting) users. Or secondly, could it be that the use of audio changes the nature of the space. In our own experience is seems as if the various locations which contain proxies are rooms next to my physical office. Just as I can overhear activity in physical rooms next to my own office, so it is with the virtual rooms. They often feel alike, as if both forms of space are in constellation around a primary physical location. It is hoped that extended use of the system will address which of these, or other factors, are crucial.

Acknowledgements

The authors would like to acknowledge the efforts of Thierry Nabeth of the Department of Technology Management at the Centre for Advanced Learning Technologies, INSEAD, France. His Java port of the AlphaWorld SDK made the Marvin system possible, and his speedy updates from our comments were invaluable.

Bibliography

Backhouse, A. & P. Drew (1992) The Design Implications of Social Interaction in a Workplace Setting. *Environment and Planning B: Planning and Design*, 19: 573-584.

Huxor, A (1999) The Role of 3D Shared Worlds in Support of Chance Encounters in CSCW. In: Vince, J. & Earnshaw, R. (eds.) *Digital Convergence: The Information Revolution.* Springer-Verlag

Towell, J. & Towell, E. (1997) Presence in Text-Based Networked Virtual Environments or "MUDS" *Presence* 6(5) 590-595.