

A Social Metaphor-based 3D Virtual Environment

Steve DiPaola
Simon Fraser University
steve@dipaola.org

David Collins
Adobe Systems Incorporated
david.collins@adobe.com

Abstract

Our design goal for OnLive Traveler was to develop a virtual community system that emulates natural social paradigms, allowing the participants to sense a tele-presence, the subjective sensation that remote users are actually co-located within a virtual space. Once this level of immersive "sense of presence" and engagement is achieved, we believe an enhanced level of socialization, learning, and communication are achievable.

OnLive Traveler is a client-server application allowing real-time synchronous communication between individuals over the Internet. The Traveler client interface presents the user with a shared virtual 3D world, in which participants are represented by avatars. The primary mode of communication is through multi-point, full duplex voice, managed by the server.

We examine a number of very specific design and implementation decisions that were made to achieve this goal within platform constraints. We also will detail some observed results gleaned from the virtual community and virtual learning user-base, which has been using Traveler for several years.

Keywords: Avatars, virtual environments, group communications

1 Introduction

Traveler is a multi-user, voice-enabled VRML browser. It employs 3D environments and avatars with complex facial animations to provide a platform for synchronous, multi-point voice communications [1,2]. The goal in developing Traveler was to deliver a rich and compelling experience of human socialization, using a common consumer PC platform over the World Wide Web. With this goal in mind, a number of very specific design and implementation decisions were made to achieve the intended level of free-form socialization, while operating within the platform constraints. These will be examined and evaluated in detail.

1.1 The Experience

In Traveler, users are immersed in a shared 3D world, with first-person perspective. Each user is able to navigate with six degrees of freedom, and each sees the other participants as fully modeled 3D characters. As a user speaks, his or her voice emanates from the corresponding avatar on each of the other clients. The avatar's lips and facial structure synchronize with the words spoken and sound of the voice is distance-attenuated and spatialized in stereo, according to its position in the 3D world relative to the local user.

The voice communication in Traveler is full-duplex and fully multi-point, i.e. the user is receiving audio while speaking and multiple streams of voice audio are delivered to each client. The overall effect of the voice delivery in conjunction with the visual environment is that of a virtual "cocktail party". Users spontaneously form and re-form conversational subgroups, using natural social conventions. Initial contacts are made using natural-world methods, such as saying "Hello, there!" towards another avatar and waiting for the other participant to turn and orient (in response to the spatialized audio cue) before proceeding.



Figure 1. Group voice chatting in Traveler via lip-sync emotive avatars.

2 Goals and Design Points

Traveler was developed in response to a number of basic observations. The first was that the developing Internet was a popular medium for multi-point chat and that real-time group communication was a swiftly growing category of application. Witness the relative success of America On-line, which concentrated on providing group chat over competitor CompuServe that focused on delivery of services. The second observation is that interleaved lines of user-typed text are a low-grade simulation of real-world social phenomena in which group communication takes place. Text-chat rooms were used to implement parties, common interest clubs, debates, discussion groups. These traditional group communication forums were simulated using what is essentially a highly artificial format that was suited to a low-bandwidth medium and that required the cognitively taxing processes of typing in real-time while simultaneously extracting multiple interleaved threads of text.

The goal in developing Traveler was to produce an intuitive communication format, which offered existing chat users a compelling experience and potential new users would find less intimidating. By including as many "organic" channels of information as the bandwidth of dial-up Internet access would allow, Traveler would allow a novice user to effectively engage in group communication by relying on intuition developed in real-world social circumstances.

2.1 Voice

The basic hypothesis in implementing Traveler was that the use of human voice is the most natural way to carry on shared conversation. The implementation of an effective multi-voice audio environment was the primary design target. This bears some emphasis, since all the other aspects of the Traveler interface,

including the 3D environment, were implemented in support of this goal. Implementing distributed voice over the Internet introduced an enormous amount of complexity to the implementation of Traveler, but it was considered essential for a number of reasons. As opposed to text chat, the use of voice leaves the hands free for use in navigating the 3D environment. The user is freed from the cognitively taxing task of extracting a stream of conversation from interleaved threads, while simultaneously typing a response. The visual focus of attention is on the other users and the non-vocal queues that expressed through their avatars, while the audio focus is on their voice. Finally, the human voice is tremendously rich in the layers of meaning expressed, beyond the simple stream words. Inflection and timing inject meaning into a sentence that is very hard to include in simple text. Witness the difficulty involved in introducing an ironic tone into an e-mail and the possibility for misinterpretation that one risks in the attempt.

2.2 Avatars and Facial Animation

In an attempt to further enhance the organic feel of the Traveler experience, the decision was made to implement avatars as smoothly morphing 3D models that animate in response to the user's voice. Usually, as in the case of anthropomorphic avatars, this animation takes the form of synchronizing the movement of the jaws and lips to the phonemes used by the speaker. This creates the profound illusion of a human face in the process of producing speech. This animation helps the user to determine which avatar in the field of view is speaking and adds to the overall illusion of being in the direct presence of living, conscious creatures. This same morphing technique is used to implement blinking, breathing, changes in emotional state and other lifelike sequences to further enhance the subtle impression of life in the avatars. Because Traveler avatars showcase the face so prominently, this organic effect is highly resonant with users, due to the extreme psychological and neurophysiological importance of the face to the human psyche. Users have reported a desire to maintain eye contact and to feel the effects of personal space during a Traveler session, indicating a high level of immersion in the social environment.

3. Community: 3D voice with 3D navigation

In designing the Traveler experience, we employed a consistent minimalism that served two primary purposes. The first was to provide a satisfying, responsive experience of an animated world on platforms with limited CPU power and communications bandwidth. The other was to keep the experience focused on a few essential channels of communication that maximized the user's sense of being actually co-located with other real individuals. Our approach was basically a narrative one. We used simple 3D graphical elements to merely suggest various elements of a world and its inhabitants, while at the same time insisting that these inhabitants are "real" by investing them with certain very organic characteristics (voice, fluid motion, emotions, autonomic twitches, etc.) The principle is not unlike the aesthetic employed in traditional animation. Highly stylized people and animals are convincingly portrayed as fully developed characters despite the fact that they bear little actual resemblance to a real-world person. We as viewers seem very ready to accept a character as a person, regardless of how fantastic their appearance as long as they have a recognizable face, are imbued with speech and follow certain familiar patterns of social behavior. By creating an environment that showcases these and other human characteristics, we endeavored to create an experience that was at the same time appropriate to the platform and uncompromising in its portrayal of a virtual place to meet real people.

3.1 Telepresence: Binding the pair

Our basic premise in creating Traveler as a social experience was that humans engage in community primarily with other humans. Thus, if a user is represented in a virtual world by an avatar, another user must perceive that avatar as a real person, if the world is to be useful as a social space. We chose to make extensive use of voice because it is such a rich, multi-layered channel of communication, which conveys a great deal of individual character. We chose to emphasize the face in our avatars because immediacy and intimacy implied by face-to-face communication. The voice belongs to the user, but is fully transferred to the avatar's face through the use of lip-syncing and virtual location. Thus we talk about the "binding the pair", the unification of the remote user and the corresponding avatar in the mind of the local viewer.

Some evidence that suggests a level of success in this binding emerged during early user tests of the system. It was observed that users felt the need to maintain eye contact with the virtual avatars on the screen. They seemed hesitant to turn away from the screen for fear of being perceived as "rude", despite being aware that their turning away could not be perceived by the other users. The suspension of disbelief in using the system was such that unconscious social patterns of behavior were in effect. Similarly, in participating in Traveler sessions, it is clear that certain standards of social behavior are naturally observed in the virtual world. Users describe a sense of discomfort when a novice user navigates too closely and thus violates the normal sense of "personal space". In response, the violated user will navigate backward to a "safe" distance. Users tend to unconsciously turn to orient on the current speaker as one would in the real world and generally organize themselves in social patterns with the virtual space. All of these things suggest a high level of immersion in the illusion of co-location.

Acknowledgements:

The authors would like to acknowledge Ali Ebtekar and Rod MacGregor, Henry Nash, Dave Owens, Stasia McGehee and James Grunke for their participation in designing and implementing OnLive Traveler. We would also like to thank the long time community members of the Traveler worlds for their support and insight.

References

- [1] DiPaola, Collins, A 3D Natural Emulation Design to Virtual Communities, Siggraph '99, 1999.
- [2] Damer, B. Avatars!: Exploring and Building Virtual Worlds on the Internet. Peachpit Press, Berkeley. 1998.
- [3] Rheingold, H. The Virtual Community: Homesteading on the Electronic Frontier. Addison-Wesley, New York. 1993.
- [4] Damer, B., S. Gold, J. de Bruin, D-J. de Bruin.. "Steps toward Learning in Virtual World Cyberspace: TheU Virtual University and BOWorld." In Interactions in Virtual Worlds. : A. Nijholt, O.A. Donk, E.M.A.G. van Dijk (eds.): University Twente, Enschede, 31-43. 1999

The full length version of this paper can be viewed or downloaded at: <http://www.dipaola.org/siggraph03>