Virtual Roommates in Ambient Telepresence Applications

Andrei Sherstyuk
University of Hawaii
andreis@hawaii.edu

Kar-Hai Chu
University of Hawaii
karhai@hawaii.edu

Sam Joseph
University of Hawaii
srjoseph@hawaii.edu

Abstract

We introduce Virtual Roommates — a mapping of loosely linked spaces, that allows users to overlay multiple physical and virtual scenes and populate them with physical and/or virtual characters. As the name implies, the Virtual Roommates concept provides continuous ambient presence for multiple disparate groups, similar to people sharing living conditions, but without the boundaries of real space.

1. Introduction

An ability to communicate is one of the most fundamental human needs. Communication technologies and applications become ubiquitous and cover many modalities: audio, video, tactile. Both the volume and nature of information exchanged vary enormously, from a single audio channel in a telephone conversation to 3D data streams from game servers that can put users into a planet-size virtual world, with geometric detail down to a single blade of grass [1].

It seems likely that virtual worlds will be playing increasingly important roles in society. Evidence for this are the numerous academic and commercial conventions on serious games, cyber-health, intelligent virtual agents and other similar topics, covering technical and social aspects of this phenomenon. That inevitably brings up the question, how one can interact with such virtual worlds and their inhabitants in a more realistic manner. The first steps have already been made: virtual characters from Second Life, a 3D online virtual world, were brought into a real environment [2]; and conversely, interactive video-based user avatars have been inserted into virtual scenes [3].

Our work pursues similar goals: provide 'cross-reality' connectivity for multiple spaces and multiple agents. In addition, we want our solution to be free from spatial constraints imposed by geometric layouts of the participating environments. Both goals can be summarized and illustrated with a single metaphor of *virtual roommates*, who coexist in loosely coupled environments that can be real, virtual, or both, as shown in Figure 1. Virtual roommates, as objects, are avatars of remote participants projected into each one's local environments. Functionally, *Virtual Roommates* is a special-purpose mapping function for localizing and visualizing users in shared environments.

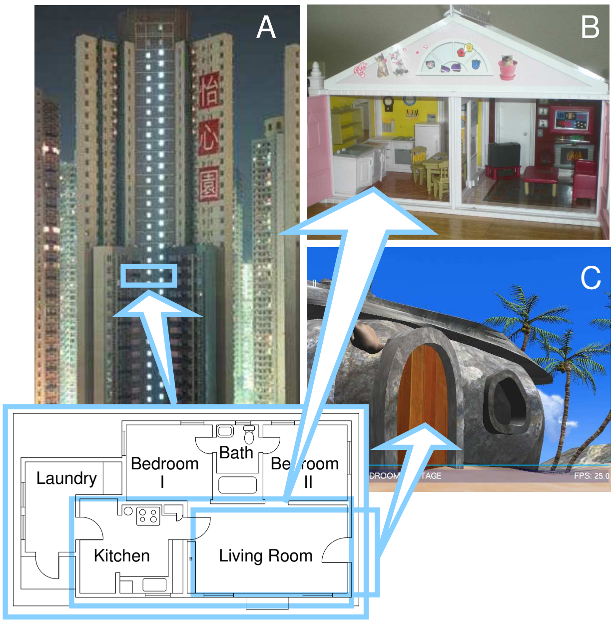

Figure 1. A real apartment in a residential building (full reality), a doll-house (make-belief reality) and a Stone Age cabin (virtuality) can be linked, populated and shared by virtual roommates.

2. Telepresense and Mixed Reality

Advances in Augmented Reality (AR) [4, 5] have made practical applications possible in different fields, particularly in collaborative systems, including telepresence and telemedicine. One example is the HyperMirror interface. HyperMirror [6] is a video conversation system that superimposes remote participants into a shared virtual video, producing the effect that they are in the same room. HyperMirror offers students the ability to see a video of their own image standing next to a remotely based instructor. This allows them to copy the instructor's stances and motions to learn and practice certain medical procedures.

ICAT 2008
Dec. 1-3, Yokohama, Japan
ISSN: 1345-1278

Another example is Cisco's Telepresence conference system, which utilizes high-definition, life size video, directional audio, and shared remote artifacts to produce an immersive effect of sharing a room. However, the feeling of immersion cannot be maintained once a person leaves the camera's field of view. This puts severe limitations on the range of activities that users can do collaboratively. Similarly, the HyperMirror interface achieves the sense of presence by creating composite images of remote users appear as if they are standing next to each other. The illusion is broken when either of the participants moves too far from their respective positions. With both the Cisco and HyperMirror applications, the effect of shared presence can only be maintained if the participants do not stray far from the camera.

We aim to remove this restriction and advance the concept of telepresence by providing persistent connectivity between all participants, from any location in their respective environments, either real or virtual.

3. Connecting virtual roommates: feature-based spatial mapping

Mixing virtual objects with a real environment involves (1) localization of the object in its own space, (2) projecting it into a destination environment and (3) visualization. In this section, we will focus on the second problem, by means of constructing a meaningful mapping between arbitrary environments.

The main principle behind *Virtual Roommates* is to abandon the dominant role of linear transformations that are commonly used for projecting objects from one location to another. Linear transformations have two important properties: they preserve collinearity of points and ratios of distances. These features are very important in applications which require precise spatial coordination between mixed spaces. Practicing medical procedures with the help from a distant instructor, linked via a HyperMirror system, is one such example.

The downside of such direct mapping is that local features of the destination environment are not taken into account. Direct mapping projects trajectories of source objects into the destination scene "as is", assuming that there is enough room to accommodate these maneuvers. This is not always true because the environments that are being mixed may have different sizes and different geometric layouts. Moreover, user actions and movements that make sense in one environment, may critically change their meaning in another. For example, a routine morning trip from the bathroom to the kitchen in a two-bedroom apartment (Figure 2, left, solid curve), may turn into a jump from a window in another apartment, if projected there verbatim (Figure 2, right, dashed curve).

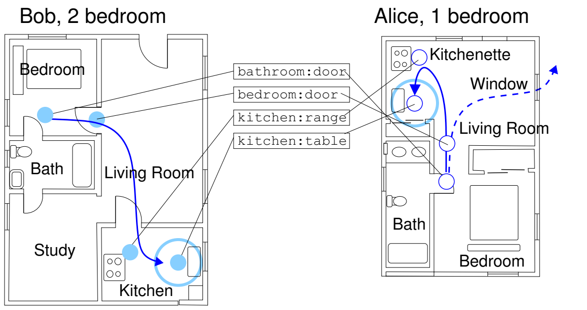

Figure 2. Mapping morning activities of two virtual roommates, Alice and Bob. Bob leaves his bathroom and heads for the kitchen (left diagram, solid curve). His path is projected into Alice's place, where she can watch his movements. They "meet" for breakfast, each in own kitchen. A dashed line shows an example of incorrect mapping of Bob's path, which makes him miss the target.

In order to project the user paths correctly, we suggest using local features of the environment as anchors for object localization. With this approach, the original user path (Figure 2, left) is described as a sequence of locations, shown as solid circles: `bathroom:door`, `bedroom:door`, `kitchen:range`, `kitchen:table`. As soon as the moving object is localized with respect to the closest anchor, this information is transmitted to the receiving side. In our example, it is a one-bedroom flat shown in Figure 2, right. There, the process is run in reverse: after receiving the `"kitchen:table"` text token, the system finds the corresponding feature in the local floor plan and derives the new object coordinates from that map.

The system tracks user position and, if possible, orientation in real or nearly real time. However, the obtained 3D coordinates are not used or shared immediately in their raw form. Instead, more descriptive albeit less frequent updates are sent to all other parties involved, which we call *sparse presence samples*. Besides user location and orientation, these samples can include elapsed time at the present position, and guessed activities.

Using such sparse sampling is a logical extension of a simple observation that in many collaborative activities, precise real-time tracking is not required. This is even more true in social situations, when people are communicating over close distances. In many cultures, it is considered rude to watch people closely and continuously. Figuratively speaking, sparse sampling provides a socially acceptable way of tracing people's movements at low rates of "few direct glances (samples) per minute". When the participants cannot see each other directly (for example, being in a different room), localization rates may also be very low.

Another argument that explicitly supports the idea of sparse sampling comes from the field of human cognition. Simons and Levin showed that people's internal representation of their surroundings is surprisingly patchy and far from being complete [7]. In other words, humans naturally sample reality at rather low rates and then fuse these samples into a single coherent model by mental extrapolation, in time and space. Therefore, adding a relatively small number of extra samples from another environment should suffice for building a continuous description of the augmented scene.

4. Technical components

The *Virtual Roommates* system has three major technical components: tracking, voice communications and visualization.

4.1. Tracking

As discussed above, the proposed system has very low spatial and temporal resolution requirements. Several tracking technologies are suitable for that purpose. One solution is offered by *Smart Floor* device, a pressure sensitive carpet fit to the floor surface of the tracked room [8]. RFID-based systems can also be used for localization [9, 10]. As described by Becker et al.[10], a person wearing a belt-mounted antenna device can be reliably localized in 20 cm range from passive RFID tags, placed in the environment. Combined tracking solutions with multiple sensor fusion are very promising, as they are able to report both user location and orientation in the scene [11, 12].

4.2. Voice communications

Voice communications between roommates must be implemented with full spatial localization. Fortunately, it is relatively easy to do for in-doors settings, using a number of speakers, installed around the room. Once the position of the remote roommate is resolved in the destination space, its voice channel must be sent to the closest speaker. The spatial layout of the local environment, such as walls and corridors, will create all necessary effects automatically, including volume attenuation, reverberation, echoes, etc.

4.3. Visualization

Visualization is the most challenging part in the process of reconstruction of user presence. The current state-of-the-art head mounted displays (HMD) are still too bulky to wear for an extended period of time. Monocle-style displays may provide a more comfortable solution. The size and weight of recent models allow them to be attached to a pair of conventional sun-glasses. For a comprehensive review of wearable display technologies we refer readers to Hainich's book [13].

We suggest an alternative solution that makes use of one or more wall-mounted monitors, operating in a 'mirror' mode. From a current user standpoint, each screen shows the interior of the room, augmented with an avatar of the remote roommate, as if reflected in a co-located mirror. In other words, virtual roommates can only be seen as reflections, but never in a direct view. Reflected views of the local environment may be obtained from live video streams or by using frequently updated photographs, for better quality. In order to resolve occlusions with the virtual characters, the system must also have an up-to-date 3D model of the scene.

We believe this configuration has potential, for the following reasons: (1) it is based on a very familiar concept: mirrors have been used for visual augmentation for thousands of years; (2) it is unobtrusive, in contrast to monocles and HMDs, and provides 'viewing on-demand': people will see their roommates only when they want to, by looking into a virtual mirror; (3) naturally solves a problem of colliding with remote avatars and other virtual elements of the augmented scene, because reflections are not tangible objects ("see only"); (4) easy to reconfigure at any time, by moving one or more screens to the location of current activity; (5) may cover large areas, depending on the current position of the viewer, as illustrated in the diagram below.

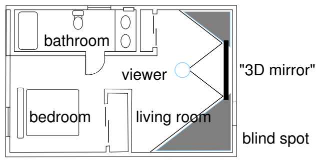

Figure 3. A single wall mounted 3D mirror allows to visualize a virtual character almost anywhere in the room.

5. Preliminary tests

In order to connect two or more real people with the *Virtual Roommates* system, all elements discussed in section 4 must be in place. Presently, we are still working on tracking, so physical testing is not possible yet. However, by using VR and miniature models, we can test nearly all aspects of our feature-based mapping, which is the core idea behind the whole system.

We took advantage of the fact that a doll house is a fully controlled tracked environment, quite suitable for building a prototype system. For that purpose, we used a little doll figurine, attached to a Flock of Bird magnetic sensor and moved her around the house, playing a simple 'dinner-then-TV-show' scenario. During that sequence, the system captured locations of the figurine, with respect to the furniture objects. Then, this sequence was replayed in a Stone Age one bedroom cottage scene, which had similar household elements. The results are shown in Figure 4. This test shows the first case of successful projection of a real-life scenario onto virtual settings with completely different geometry.

We also created a prototype of a virtual mirror, described in section 4.3. The results are shown in Figure 5.

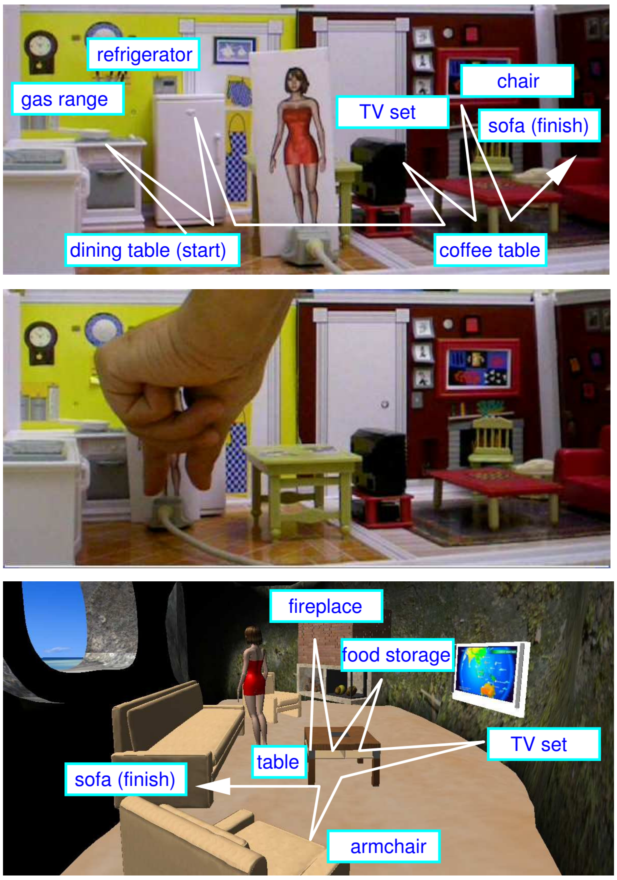

Figure 4. Tracking location of a toy character inside the doll house using furniture items as landmarks. Locations and orientations of the figurine are then reconstructed inside the virtual Stone Age cabin, with similar elements. For external views, see Figure 1.

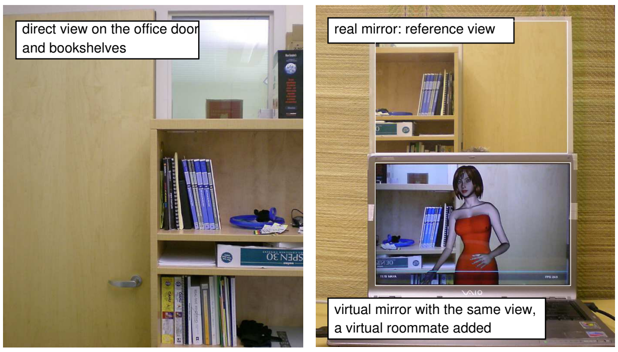

Figure 5. Virtual mirror prototype: adding a virtual roommate to an office scene. A real mirror is mounted on top of a laptop, providing a reference view of the office. The laptop operates in a virtual mirror mode, showing the same scene plus an animated virtual character.

6. Conclusions

The borderline between real and virtual worlds is getting increasingly fuzzy. People live their alternative lives in VR: make friends and enemies, express themselves creatively and socially, as well as making a living. The *Virtual Roommates* system, outlined in this paper, describes the very first steps towards a true fusion between people's real and virtual lives.

References

[1] Blue Mars Project: Massively Multiplayer Virtual World. *Official Press Realease*, http://www.avatar-reality.com/index-6.html, Feb 19, 2008. 1

[2] T. Lang, B. MacIntyre, I. Zugaza. Massively Multiplayer Online Worlds as a Platform for Augmented Reality Experiences. *Proceedings of IEEE Virtual Reality Conference*, Reno, Nevada, pp. 67-70, 2008. 1

[3] S.-Y. Lee, S. Ahn, M.-T. Lim, H.-G. Kim. Tangible Video Avatar for Natural Tele-Interaction. *Proceedings of 4th International Workshop on the Tangible Space Initiative, ISMAR*, Nara, Japan, 2007. 1

[4] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, B. MacIntyre. Recent Advances in Augmented Reality. *Computer Graphics & Applications*, 21(6), pp. 34-47, 2001. 1

[5] G. Papagiannakis, G. Singh, N. Magnenat-Thalmann. A survey of mobile and wireless technologies for augmented reality systems. *Computer Animation and Virtual Worlds*, 19, pp.3-22, 2008. 1

[6] O. Morikawa, T. Maesako. HyperMirror: toward pleasant-to-use video mediated communication system. *Proceedings of the ACM Conference on Computer Supported Cooperative Work*. Seattle, Washington, Nov. 14-18, 1998. 1

[7] D. Simons and D. Levin. Change Blindness. *Trends in Cognitive Sciences*, 1(7), pp. 261-267, 1997. 3

[8] R. Orr, G. Abowd. The Smart Floor: A Mechanism for Natural User Identification and Tracking. *Conference on Human Factors in Computing Systems*, Netherlands, April 2000. 3

[9] P. Bahl, V. Padmanabhan. RADAR: An in-building user location and tracking system. *Proceedings of the IEEE INFOCOM 2000*, pp. 775-784, 2000. 3

[10] B. Becker, M. Huber, G. Klinker. Utilizing RFIDs for Location Aware Computing. *UIC, Lecture Notes in Computer Science*, Springer, vol. 5061, pp. 216-228, 2008. 3

[11] D. Hallaway, T. Höllerer, S Feiner. Bridging the gaps: Hybrid tracking for adaptive mobile augmented reality. *Applied Artificial Intelligence, Special Edition on Artificial Intelligence in Mobile Systems*, 25(5), July 2004. 3

[12] D. Pustka, G. Klinker. Dynamic Gyroscope Fusion in Ubiquitous Tracking Environments. *Proceedings of the 7th IEEE International Symposium on Mixed and Augmented Reality, ISMAR*, pp. 13-20, Cambridge, UK, 2008. 3

[13] R. Hainich. *The End of Hardware: a Novel Approach to Augmented Reality (2nd Ed.)*. BookSurge Publishing. 2006. 3