Virtual Roommates in Multiple Shared Spaces

Andrei Sherstyuk*
Avatar Reality, Inc.

Marina Gavrilova †
University of Calgary

ABSTRACT

Augmented Reality applications have already become a part of everyday life, bringing virtual 3D objects into real life scenes. In this paper, we introduce "Virtual Roommates", a system that employs AR techniques to share people's presence, projected from remote locations.

Virtual Roommates is a feature-based mapping between loosely linked spaces. It allows to overlay multiple physical and virtual scenes and populate them with physical 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.

Index Terms: H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities— Evaluation/methodology; I.3.7 [Computer Graphics]: Three-dimensional Graphics and Realism — Virtual Reality ; I.3.6 [Computer Graphics]: Methodology and Techniques — Interaction techniques

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 volume and nature of information exchanged range immensely. Examples include: a telephone single voice channel; a video stream between a doctor and a patient, with embedded virtual content; a tele-conference system, capable of connecting dozens of people in high definition video. Online virtual worlds present a very special case of information exchange between communicating parties. Game servers provide data for planet-size worlds, with geometric details to a single blade of grass. The amount of 3D content available online, now exceeds what can be explored in a single person's lifetime.

It seems likely that online virtual worlds will be playing an increasingly important role in society. Evidence for this are numerous academic and commercial conventions on serious games, cyberhealth, intelligent virtual agents and multitude of similar topics, covering technical and social aspects of this phenomenon. Assuming that this trend will continue, the following questions come to the front. How one can interact with such virtual worlds and their inhabitants in a more realistic manner? Can virtual worlds be used to enrich person-to-person communications in everyday life?

In the past few years, a number of projects demonstrated examples of meaningful interaction between real and virtual environments: Human Pacman Augmented Reality (AR) game [7], AR Façade [8], AR Second Life [15]. In these projects, life-size virtual characters were brought to real-life environments, both indoors and outdoors. Conversely, techniques were developed that allowed to insert user avatars into virtual [13] and real scenes [17]. One of the latest results in this area is reported by Spanlang and colleagues [30], who built a realistic avatar of a real person, with full body tracking and collision detection, ready for integration into VR and AR applications.

*e-mail: andrei@avatar-reality.com
†e-mail:mgavrilo@ucalgary.ca

The research presented below pursues similar goals: provide 'cross-reality' connectivity for multiple spaces and multiple agents. In addition, the solution is sought to be free from spatial constraints imposed by geometric layouts of 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 mixed.

Virtual roommates, as visible entities, are 3D avatars of remote participants projected and visualized in each one's local environment. Functionally, each "roommate" is also an instance of a special-purpose function set for sampling and reconstructing presence of mobile agents in shared spaces. As a whole, the Virtual Roommates system is an extension of conventional telepresence applications into the AR domain, as was recently suggested by Sherstyuk and colleagues [27, 28]. In this work, we extend and further develop the feature-based spatial mapping between linked spaces, which constitutes the core principle of the Virtual Roommates system.

Before presenting the architecture of Virtual Roommates, we will examine several existing systems, capable of connecting both real and virtual environments and content. We will demonstrate that despite their seeming flexibility, most telepresence applications put tight restrictions on the geometry of connected spaces and allowed user activities. Removing these restrictions constitutes the goals of this paper.

2 BACKGROUND: FROM TELE-COLLABORATION TO AUGMENTED REALITY COEXISTENCE

Advances in display technologies and growing availability of broadband networks spawned a number of commercial and open-source teleconference systems. The *TelePresence* line of products, launched in 2006 by Cisco Inc., represents the first group [31]. In a TelePresence installation, two dedicated rooms are connected into one shared teleconference environment, using HDTV displays and spatialized sound. TelePresence can accommodate up to eighteen people on the "local" side of a conference table. More examples of multi-display collaborative systems may be found in a survey by Ni et al. [22].

The *Access Grid* is an example of a non-commercial system facilitating group-to-group collaboration. It was developed in the USA at Argonne National Laboratory and first deployed by the National Center for Supercomputing Applications in 1999. Access Grid (AG) can connect multiple rooms in arbitrary locations, provided that all participating sites are registered as AG-nodes. During conference sessions, each connected node is represented by its own window, usually projected on a large screen in tiled mode. Because of its flexibility and relatively low operation costs, Access Grid become popular in academic community. Overall, the advent of large displays and broadband networking made tele-collaboration not only possible, but also practical and cost effective [22].

Concurrently, advances in microdisplay and motion tracking technologies made Virtual Reality components available to a wider audience of researchers. The increased availability of VR equipment has stimulated creation of novel interfaces and applications, especially in the Augmented Reality field, where virtual content is superimposed onto real scenes. An overview article by Azuma et al. [1] gives a historic perspective on the scope of traditional AR applications. A recent survey by Papagiannakis, Singh and Magnenat-Thalmann describes advances in mobile and wearable AR technology [25]. A comprehensive literature review by Yu et al. [33] gives a very broad view on the current state of AR, with over three hundred references. Below, we will examine several projects, directly related to collaboration between remote participants and cross-platform communications.

One of them is the HyperMirror telecommunication interface, developed by Morikawa and Maesako [20]. HyperMirror is a video conversation system that superimposes mirrored images of remote participants onto a shared video stream, producing the effect that they are standing side-by-side with the viewer. Taking advantage of modern tracking technologies, the HyperMirror technique was recently used in a distance education system for teaching endoscopic surgery skills, developed by Kumagai et al [14]. Using this system, students are able to watch themselves, standing next to a remotely based instructor. By copying instructor's stances and motions, students learn and practice a variety of surgical procedures.

Another example of employing augmented video for training purposes is described by Motokawa and Saito [21], who created a system for teaching basic guitar skills. Their system captures player's hands on video and overlays it with the rendered 3D model of virtual hands, demonstrating guitar chords. To learn the chords, the player simply overlaps his own fingers on the virtual hand, looking at the mirrored image on the monitor. In this example, real players collaborate with the virtual instructor, using shared working space over the real guitar object.

AR Second Life and *AR Façade* projects, developed at Georgia Institute of Technology, both feature immersive Augmented Reality settings suitable for modeling social situations that involve virtual characters. In AR Second Life project, a resident of a virtual world *Second Life* is projected and visualized in a real lab environment [15]. In AR Façade installation, participants interact with a virtual married couple, while moving freely inside a physical apartment [8]. They get engaged in a conversation with the virtual husband and wife, using natural speech and gestures. To avoid errors introduced by automatic speech recognition, user verbal input is manually typed-in in real time by a human operator behind the scene, using a Wizard-of-Oz approach. The virtual couple also move around the same apartment. Their life-size avatars are visualized with a Head Mounted Display.

Despite the diversity of these projects and systems, all of them require the collaborative spaces to be identical geometrically, for all participants involved, real and virtual. In teleconference systems, such as TelePresence [31], the local and remote spaces are usually tiled back-to-back, connected by the surface of display screens. In mirror-based systems [14, 21], the virtual or remote environment is overlaid onto a local scene, using reflection transformation. When immersive visual augmentation is used [8, 15], the virtual and physical spaces must coincide precisely. In addition, in video-based systems [31, 14, 21] the feeling of shared presence can only be maintained while a person remains in the camera's field of view. The illusion is broken when participants move too far from their dedicated positions.

The described restrictions put severe limitations on types of environments that can be linked and shared and also on range of activities that users can do collaboratively. In the next section, a non-linear mapping between linked spaces will be introduced, which will provide geometry-free persistent connectivity between all participants.

3 Connecting Virtual Roommates via Feature-Based Spatial Mapping

Sharing objects from different environments requires: (1) localization of the object in its own environment; (2) projecting it into the destination space and (3) visualization of the object.

In this section, the second problem will be addressed, by constructing a meaningful mappings between arbitrary environments, that can be shared by virtual roommates. Without the loss of generality, we will illustrate the main concepts using two fictional characters, Alice and Bob, who reside in physically separate apartments, as shown in Figure 1, and use the Virtual Roommates system for spending time together. For clarity, it is assumed that each participant can be localized in their own space, by using real-time tracking. Also, it is assumed that each participant has all required equipment for visualizing the remote party. In Section 4, several solutions will be discussed for tracking and visualization.

The main idea behind the proposed system is to dismiss linear transformations as the means of projecting user paths from one environment to another. Linear transformations have two important properties: they preserve collinearity of points and ratios of distances. These features are very important in applications that require precise spatial coordination between connected or mixed spaces, for purposes of monitoring or manipulating objects. Practicing medical procedures with the help from a distant instructor, linked via a HyperMirror system, provides one such example [14].

The downside of direct mapping is that local features of the destination environment are not taken into account. Direct mapping, performed via linear transformations, projects original trajectories of 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.

Figures 1 and 2 illustrate the problem. Bob lives in a small one bedroom apartment. He has a virtual roommate Alice, who lives in a large penthouse in other city. As a part of his morning routine, Bob leaves the bedroom and moves towards the kitchen for breakfast (see Figure 1, left diagram). Alice is waiting for him at the dining table, in her place (right diagram). However, direct projection of Bob's original path onto Alice's room makes him appear to leave the bedroom and then immediately exit the room (right diagram, dashes curve).

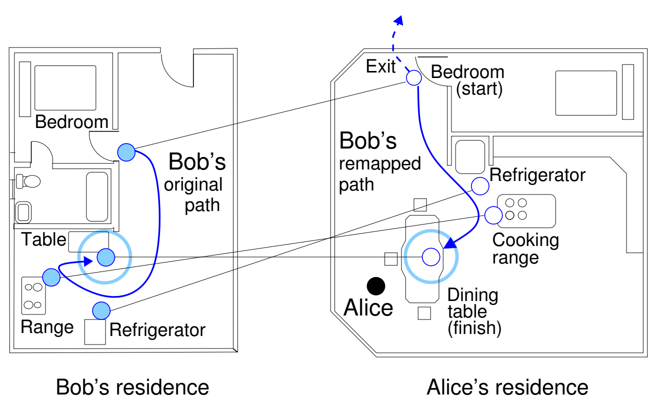

Figure 1: Morning activities of two virtual roommates. Bob leaves his bedroom and heads for the kitchen (left). His projected avatar is displayed at Alice's place, where she waits for him by the table (right, solid curve). A dashed line shows the direct mapping of Bob's path, which makes him miss the destination.

In order to project paths correctly, we suggest using local features of the environment as anchors for object localization. With this approach, the original user path (Figure 1, left diagram), is described as a sequence of locations, shown as circles:

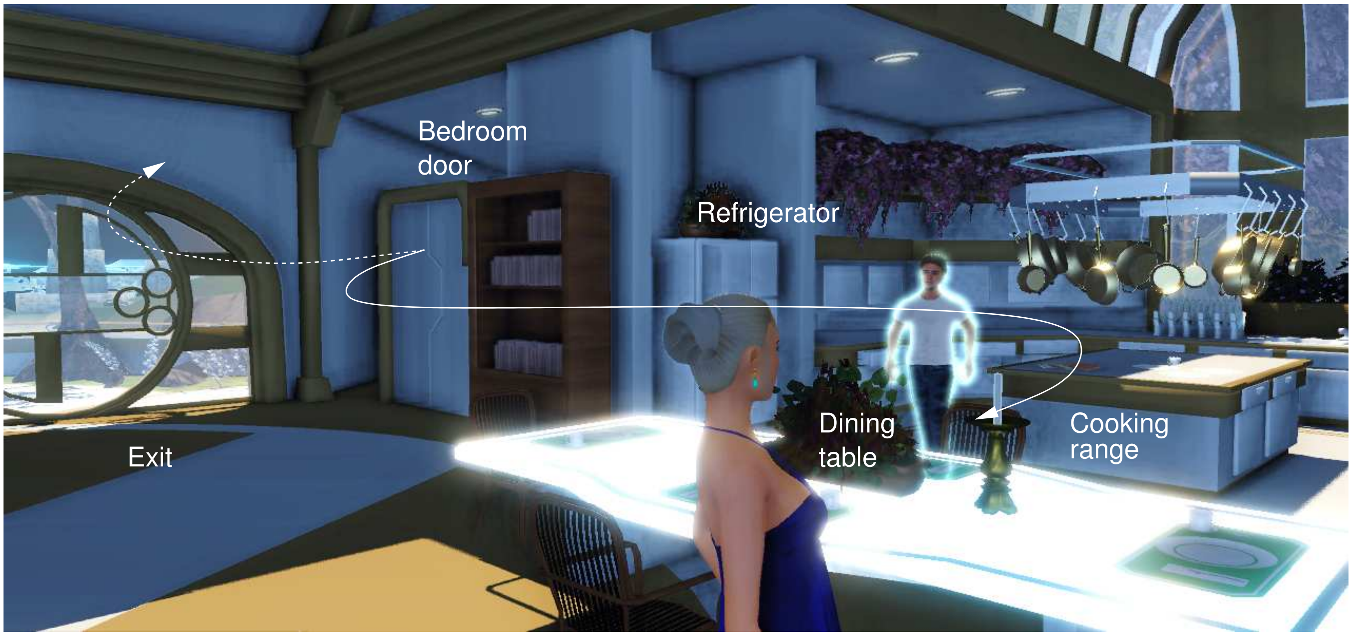

Figure 2: 3D reconstruction of the scene, outlined in Figure 1 and described in Section 3. A solid curve shows the correctly resolved path, leading towards the dining table. This illustration shows an augmented scene, where Alice and her room are real and Bob's avatar is virtual. The contour of Bob's avatar is high-lighted, to indicate its virtual nature. The scene was staged and rendered in Blue Mars platform [5].

```
"bedroom door",
"refrigerator",
"cooking range",
"dining table"
```

As soon as Bob is localized with respect to the closest anchor, this information is transmitted to the receiving end (Alice's place), where the process is run in reverse. For example, when the system receives a "dining table" anchor information, it finds the corresponding feature in the local floor plan and obtains the new coordinates from that map. The new reconstructed path is now based on local features of the destination environment. It leads Bob to his intended target, the dining table, via a series of moves from one anchor to another. The reconstructed path is shown in Figure 1 (right diagram, solid curve), and that is how Alice will see movements of Bob's avatar, projected into her penthouse.

To summarize: the Virtual Roommates system tracks user position and, if possible, orientation, in real or nearly real time. 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 can be called *sparse presence samples*. These samples include code names of the local features in the source environment (e.g., dining table, refrigerator, etc.). Every received *presence sample* is then reconstructed into local 3D coordinates in the destination space, according to the local floor plan. Besides user location and orientation, a presence sample may include additional information, such as elapsed time at each location, and guessed activity. One example of a detailed presence sample is shown below:

```
location:          refrigerator,
orientation:       12 o'clock (head-on)
elapsed time:      20 seconds
guessed activity: accessing content
```

To put it more formally, in our context *a presence-sample* denotes a contextually enriched spatio-temporal data record. By contextually enriched we understand that the preceding or current actions might indicate the intention of a person at that location; spatio means in real or virtual space, and temporal indicates that the time is an added attribute. Furthermore, presence samples can be categorized as strong and weak, depending by the amount of information they convey and/or probability that the guessed activity is correct. The example with the refrigerator above is an example of a strong sample. For convenience, we will also use the term "anchor" when referring to spatial attribute of the presence sample, meaning that we are only considering its location.

3.1 Resolving and Optimizing User Path in Destination Space

As illustrated in Figures 1 and 2, Bob's avatar completely misses the travel target if he mimics his protagonist's moves and turns verbatim: leave the bedroom, turn right, go straight. This path will lead him to an unexpected exit. Feature-based navigation will help to avoid such mistakes. However, simple jumps from one anchor location to the next is not acceptable, in the general case, and the navigation system must deal with at least two issues: (1) obstacle avoidance and (2) path disambiguation and optimization.

A brute force solution for the above task would include introducing a high number of dummy anchor locations at regular intervals, covering the whole floor plan, and searching for the shortest path, connecting point A and point B, traversing those anchors. Even for relatively small and confined indoor scenes, this solution might require processing of n^2 anchors, and the search in this space for the shortest path would involve traversing the graph with n^4 edges. Here, n is a number of distinct positions that a person can take while crossing the scene in any direction, in destination space. The room length, measured in steps, may be used as a simple estimate for n. The computational complexity of the path planning process in this case may become prohibitively high.

Luckily, there exist approaches that can be utilized to resolve the problem more elegantly and with less computational resources required. To do this properly, one must resort to computational geometry methods recently introduced in the area of mobile navigation and robotics. Solutions to the path planning problem differ fundamentally, depending on the space representation, the task at hand and the connectivity model of the space. The recent article by Bhattacharya and Gavrilova [4] classifies the main approaches to path planning as: (a) the roadmap method, (b) the cell decomposition method and (c) the potential field method. The roadmap method captures the connectivity of the free space using curves or straight lines. The cell decomposition method partitions free space using grid cells so that the edges of the cells represent the connectivity. The potential field approach fills the free area with a potential field in which a mobile agent (for instance, a robot) is attracted towards its goal position while being repelled from obstacles. Among those methods, the roadmap approach utilizes different kinds of graphs to represent the free space connectivity. Some of the common ones are probabilistic roadmaps which are graphs with vertices that are randomly generated in free space, visibility graph in which the graph vertices are the vertices of the obstacles themselves and Voronoi diagrams, whose edges represent the maximum clearance path among the obstacles. Figure 3 shows two types of roadmaps based on visibility graph (left) and Voronoi diagram (right). While the visibility graph based approach creates a complex of inter-connected paths in the environment with triangular-shaped obstacles, the Voronoi diagram based method provides a solution with clearly identified path that can be easily followed by a virtual roommate.

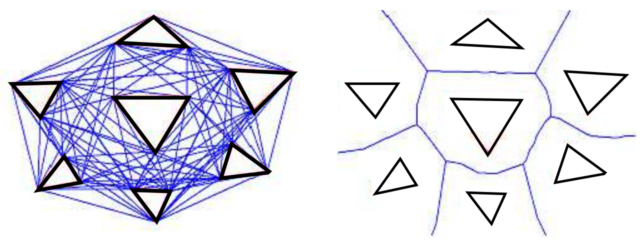

Figure 3: Left: roadmap from the visibility graph (includes obstacle edges). Right: roadmap from the Voronoi diagram.

In order to generate user paths in the destination space correctly and efficiently, we propose to combine the roadmap based Voronoi diagram method for space subdivision with the use of local features of the environment as anchors for object localization.

With this approach, the original user path (Figure 1, left) is described as a sequence of locations: bedroom door, refrigerator, cooking range, kitchen table. In the destination space, these locations must exist in the local map, used for rebuilding the path withing the new environment. The map is represented as the Voronoi diagram, with sites based on local anchor positions. Initially, the direct roadmap method is applied to quickly schedule a collision-free path, but this path is by no means optimal. It may have unnecessary turns and might not be the shortest one. This necessitates some kind of refinement on the path to achieve optimality. We propose to use heuristic based on the Voronoi diagram to obtain the optimal path. Applying a shortest path algorithm on the initial roadmap refines the path iteratively until an optimal path is obtained. The method is guaranteed to converge after $n\ log(n)$ steps, where n is the number of anchors. This approach has many advantages aside from fast optimal path planning. For instance, if the topology of underlying space changes, the method can adapt to the new room layout very quickly, without the need to create a new set of dummy features or a new space grid. Also, if we now have to plan the path for a larger moving entity (i.e. a different user or a virtual object with bigger dimensions), we can simply set a new parameter to a proportionately larger value and the heuristic will still find an optimal path.

3.2 Sparse Presence Samples, Additional Remarks

Besides providing a convenient basis for computationally efficient resolution of user paths, the idea of sparse sampling has many other advantages.

Use of sparse sampling follows naturally from an 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 at close distances. In many cultures, it is considered impolite and even 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 different rooms), the sampling rates may be even lower.

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 [29]. In other words, humans naturally sample reality at rather low rates and then fuse these samples into a single coherent model by mental extrapolation, both in time and space. Therefore, adding a relatively small number of extra visual samples from another environment should suffice for building a continuous description of the augmented scene. Kim and colleagues first suggested [12] and then demonstrated experimentally [16], that people are able to maintain tight association with virtual elements in augmented scenes, without tight spatial and visual registration.

4 System Components Required for Sampling and Reconstructing User Presence

To explain how the system captures and projects user paths from one environment to another, we assumed that both virtual roommates, Alice and Bob, have sufficient means to track their physical movements and visualize themselves as 3D avatars in each other's local environment. In this section, we will discuss technical details and evaluate several possible implementations.

4.1 Tracking

It was shown, that spatial and temporal tracking resolution requirements can be very low during path sampling, when the target object (Bob) is localized and resolved in the destination space, Alice's apartment. A relatively small number of anchors is sufficient for path planning, including obstacle avoidance and path optimization. However, direct immersive visualization of a remote roommate requires, in general, precise registration of its avatar with a local viewer, which calls for real-time tracking of the viewer's head, in six degrees of freedom (DOF).

General-purpose 6 DOF tracking can be achieved using commercially available high-end systems, such as PPT from WorldViz [1]. PPT is able to capture object position and orientation in a 50 x 50 x 3 meter volume, with sub-millimeter precision. PPT tracker is using multiple infrared cameras that entirely cover the working area. The user must wear a head mounted IR emitter, which is reasonably comfortable and unobtrusive. The only serious issue with this solution is the price of the tracking system.

Combining tracking technologies with multiple sensor fusion provides an alternative solution to the 6 DOF tracking problem [11, 26]. For example, in a system developed by Tenmoku and colleagues [32], five wearable sensors are used: a head-mounted camera, an inertial InterTrax2 sensor, an IrDA receiver, an RFID tag reader and a pedometer. Combined input from all these devices enables accurate registration of virtual content for current user position and orientation, both in indoors and outdoors scenes. Tracking systems that employ sensor fusion are generally more flexible and

[1] Precision Position Tracker. http://www.worldviz.com/products/ppt

customizable than off-the-shelf systems and allow building less expensive applications. However, use of multiple wearable sensors often results in very bulky and uncomfortable sets on interconnected components, including a laptop, PC or ultra-mobile PC, mounted on a special harness.

Presently, video-based tracking technologies experience rapid growth, partly energized by commercial success of EyeToy line of video games by Sony. In these games, players control their video avatars, using natural body movements, tracked by a TV-mounted camera. Video-based tracking can be used to obtain object position and orientation, in real time. Using adaptive view-base appearance model, Morency and colleagues developed a system for full 6 DOF tracking of user head [19], with rotational accuracy and resolution of InertiaCube2 sensor from InterSense Inc.,[2] which became a de-facto industry standard in capturing rotation. This technique was demonstrated to be sufficient for real-time tracking and registration of user head at 75 Hz, running on a modest consumer level PC. The main advantage of video-based tracking is its unobtrusiveness: users are not required to wear any devices. The downside is relatively small working range, limited by camera field of view. In order to cover a reasonable working volume (e.g., a living room), multiple cameras are likely to be installed.

In order to build a practical and cost-effective tracking system for Virtual Roommates, we suggest to take advantage of the discrete nature of the tracked space. As discussed before, sampling and reconstruction of user paths assumes and requires that the working volume is partitioned into a set of local features. We suggest to use these features for initial localization of roommates, which can be further refined using video-trackers, installed at these locations.

For initial coarse localization, several techniques may be used. An interesting tracking solution is offered by *Smart Floor*, a pressure sensitive carpet fit to the floor surface of the tracked room. This device showed good results for indoors environments [24]. However, it has one serious drawback: tracking stops as soon as a person looses contact with the carpet surface, for example, when sitting on a chair or a sofa.

Radio-frequency identification (RFID) can also be used for tracking a person's location within large areas [2, 3]. As reported by Becker et al. [3], a person wearing a belt-mounted antenna device can be reliably localized in 20 cm range from passive RFID tags, placed in the environment. Their system was successfully tested in a real-life situation, namely, an aircraft maintenance routine. A crew member was detected at every passenger seat in the aircraft cabin. This resolution is quite sufficient for our purposes.

RFID-based approach fits closely to our needs and purposes. Tags can be easily attached to all feature elements in the room, and marked as "kitchen table", "lounge sofa", et cetera. RFID tags are very durable and inexpensive, which makes the whole system scalable and easy to install and maintain. Wearable RFID readers are also getting better and smaller; several commercial and home-made devices are described by Medynskiy et al [18].

A summary: for Virtual Roommate tracking system, we propose a two-stage approach. First, a person is localized in the environment, using proximity-based techniques, such as RFID, with respect to local features. These sparse positional samples are used for projecting and reconstructing the person's path in remote environment. Also, these samples are subject for further refinement, using locally installed video trackers. The number and density of placements of proximity tags and video-cameras will depend on the size and layout of the physical space and the choice of visualization method. Several visualization options are discussed in the next section.

[2] InterSense Inc. Inertia Cube2 Manual. http://www.intersense.com.

4.2 Visualization with Wearable Display Devices

Visualization is the most challenging part of the process of reconstruction of user presence. Most commercially available head mounted displays (HMD) are still too bulky and heavy to wear for an extended period of time [6]. Monocle-style displays manufactured by Creative Display Systems and Tek Gear may offer a more comfortable alternative, provided that wireless configurations will eventually appear on the market. The size and weigh of recent near-eye monocular displays allow to attach them onto a pair of conventional sun-glasses. A comprehensive review of available wearable headset and discussion of trends in display technologies may be founds in Hainich's book [10].

Recently, Sensics Inc, a manufacturer of high-end panoramic HMDs, announced a new line of products that will provide a low-latency high definition wireless video link, compatible with all Sensics' headset models. Their goal is to achieve HD1080p video streaming to battery operated HMDs at 60Hz frame rate, including the ultra-portable xSight HMD model. This wireless configuration may provide a feasible visualization option for Virtual Roommates, although the price of the whole system may become prohibitive, even for a pilot project, with a single-viewer installation.

4.3 Virtual Mirror

Until light-weight wireless Head Mounted Displays become available and affordable, we suggest an alternative indirect solution for visualization of 3D avatars, utilizing the concept of a virtual mirror. The display system will contain one or more wall-mounted monitors, operating in a 'mirror' mode. Each monitor will show the interior of the local room, where the viewer is presently located, augmented with rendering of the avatar of a remote roommate. The scene must be rendered as if reflected in a mirror, co-located with the display, and viewed from the current standpoint of the local observer. In other words, virtual roommates can only be seen as reflections in mirrors, but never in the direct view. Reflected views of the local environment may be obtained from live video streams or by using photographs of the room. In order to correctly resolve occlusions with the virtual characters, the rendering system must also have a current 3D geometry model of the local scene. However, in the simplest configuration, the local environment may be approximated by a single cube, textured with high-resolution photographs.

One important feature of Virtual Mirror as a display solution for our system is that it may eliminate the need for full 6 DOF tracking of user head. Switching from 6 DOF to 3 DOF tracking will significantly reduce requirements to system hardware components. In order to render the scene in reflected mode, it is sufficient to obtain position of the viewer in the destination environment. Also, the system must know the current location and orientation of the mirror display. That information is sufficient to render the local environment from the local user position, using standard off-axis rendering technique. The user head rotation will only be needed if stereo-rendering is required and proper stereo-separation of both eyes must be updated, with respect to the mirror location, in real time.

In addition, the virtual mirror solution has a number of attractive features, that are very relevant for the purposes of this project:

- Virtual Mirror is based on a very familiar device: mirrors have been used for visual augmentation for thousands of years, therefore, no training or suspension of disbelief is needed;

- The display device is unobtrusive, in contrast to monocles and HMDs;

- Provides "viewing on-demand": people will see their roommates only when they want to, by looking into a virtual mirror;

- Naturally solves the problem of colliding with remote avatars, because reflections are not perceived as tangible objects, they are “see only”;
- Is easy to reconfigure at any time, by moving one or more screens to the location of current activities, similarly to a dynamic visualization system, recently presented by Ohta [23].

Finally, visualization even with a single virtual mirror screen, may provide a large viewing area, depending on the current position of the viewer. In Figure 4, the shaded part of the diagram indicates the area where Alice can see her friend while standing close to the mirror. This area covers most of the room.

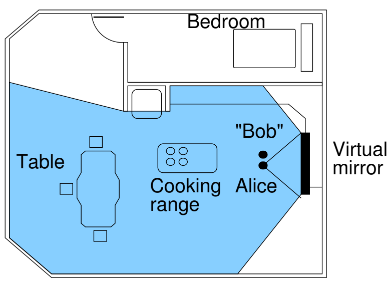

Figure 4: Room visibility diagram, for the current standpoint. See also 3D scene reconstruction in Figure 5.

Figure 5: Visualizing roommates as reflections in a virtual mirror. Top: Alice is looking into a 3D mirror, mounted above the kitchen sink. Bob’s avatar is shown as a silhouette, not visible in direct view. The bottom image shows what Alice sees in the mirror, including fully visible Bob’s avatar. Images captured in Blue Mars [5].

5 Building Prototypes in Virtual Reality

In the previous sections, several solutions were discussed for tracking and visualization of connected remote participants. A fully functional Virtual Roommates system requires all these components developed, assembled and installed at two or more physical locations, which still presents a significant challenge. However, by using VR and miniature models, it is possible to implement and test all, or nearly all, aspects of the proposed feature-based spatial mapping that constitutes the core principle of the Virtual Roommates.

5.1 Path Projection Test: Tracking Figurines in a Doll House

The prototype system, described next, took advantage of the fact that a doll house is a perfect example of fully controlled environment, where tracking can be easily implemented. A doll house also has all elements, required for building a feature-based spatial mapping. These elements are miniature furniture models (TV set, table, sofa, etc), that can be rearranged inside the house as desired. For a virtual roommate, a little figurine was used, attached to a Flock of Bird magnetic sensor, for tracking. The doll house and the figurine constituted the source space. The destination space was implemented in a virtual environment, with very different content: a Stone Age village, with one cabin complete with similar furniture items. The room layouts of the doll house and the cabin, shown in Figure 6, were completely different.

Figure 6: Mixed reality prototype of the Virtual Roommates. The little paper figurine moves around the doll house, and her locations are mapped and reconstructed inside a virtual Stone Age cabin. Both places are loosely connected via similar elements in their interiors. See also Figure 7.

During the test, an experimenter was playing with the figurine, “cooking dinner” and “watching TV” inside the doll house, while monitoring evolutions of her 3D avatar in the destination scene, the virtual cabin. The avatar mirrored these actions, moving in the cabin, from one landmark to the next. Two processed were running concurrently, sampling and reconstruction. The sampling process continuously read data from the magnetic tracker and checked for collisions between the geometry model of the figurine and the models of all furniture items. Upon detecting a collision, the name of the colliding object was sent to the reconstruction process, which controlled movements of the avatar in VR. Every new feature elements was checked with the local floor plan, and, if the corresponding item was found, the avatar moved towards the new location. As the result, a meaningful feature based path projection was established and maintained between physical and virtual environments, with different spatial layouts.

5.2 Visualization Test: A Virtual Mirror Prototype

The virtual mirror prototype was created and tested as a a standalone VR application. A real office room was approximated in 3D, by a number of boxes, textured with high-resolution photographs. Then, a virtual character was added to the virtual office scene, as shown in Figure 8. The combined scene was rendered on a laptop in a mirror mode, using the laptop screen as a display device. For a reference view, a real physical mirror was attached to the laptop.

Both mirrors showed the same room with nearly identical fidelity. The virtual mirror also showed a virtual roommate, moving around the office. For this test, no collision detection was implemented.

Both prototypes were developed using Flatland, an open source VR system, developed at the Albuquerque High Performance Computing Center [9].

6 Conclusions

The borderline between real and virtual worlds is getting increasingly fuzzy. People begin to live their alternative lives in VR: make friends and enemies, express themselves creatively and socially, and even earn real money in virtual economies. On the other hand, advances in miniaturization of tracking and display devices allow building systems that effectively insert real people's presence into virtual environments and, reciprocally, bring 3D characters into physical scenes. Further development of online virtual worlds, coupled with continuously improving input/output interface devices will call for new conceptual solutions, aiming to achieve a true fusion between people's real and virtual existence.

The Virtual Roommates system, presented in this paper, describes the first steps towards this goal. It provides a complete framework for implementing continuous ambient presence for multiple disparate groups, capable of connecting both real and virtual environments and context. It presents a number of innovations in the way tele-collaboration is integrated with the augmented reality, using the feature-based spatial mapping of virtual roommates' paths. It also provides detailed description of system components, including tracking technologies, wearable display devices and a novel application of the virtual mirror concept. The preliminary implementation of described system strongly supports the introduced concepts.

References

[1] 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.

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

[3] 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.

[4] P. Bhattacharya, M. Gavrilova. Roadmap-based Path Planning – Using the Voronoi Diagram for a Clearance-Based Shortest Path. *IEEE Robotics & Automation Magazine (IEEE RAM)*, Special Issue on Computational Geometry in Robotics. 15(2): 58-66, 2008.

[5] *Blue Mars Virtual World Platform*, by Avatar-Reality Inc., http://www.bluemars.com.

[6] Christoph Bungert. HMD Comparison Chart. http://www.stereo3d.com/hmd.htm, last updated on Dec. 27, 2006.

[7] A. Cheok, S. Fong, K. Goh, X. Yang, W. Liu, F. Farzbiz. Human Pacman: a sensing-based mobile entertainment system with ubiquitous computing and tangible interaction. *Proceedings of Workshop on Network and system support for games*, ACM, New York, NY, USA, 106–117, 2003.

[8] S. Dow, B. MacIntyre, M. Mateas. Styles of play in immersive and interactive story: case studies from a gallery installation of AR Façade. *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, ACE'08*, Yokohama, Japan, pp. 373-380, 2008.

[9] *Flatland VR engine*. The Homunculus Project, Albuquerque High Performance Computing Center (AHPCC), http://www.hpc.unm.edu/homunculus, 2002.

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

[11] D. Hallaway, T. Höllerer, S Feiner. Bridging the gaps: Hybrid tracking for adaptive mobile augmented reality. *Applied Artificial Intelligence,*

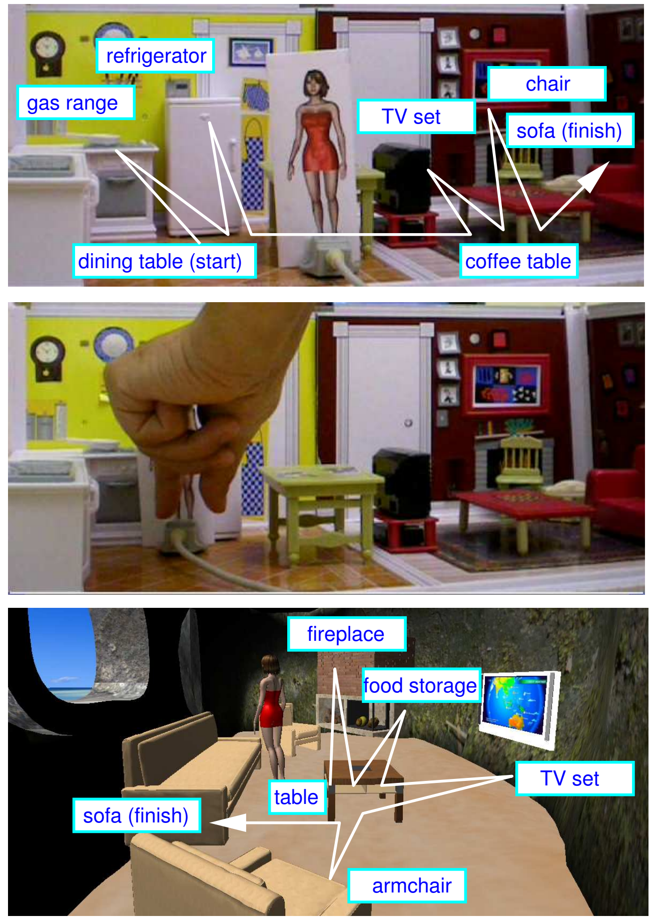

Figure 7: Tracking locations of a toy character inside the doll house, using furniture items as landmarks. Locations and orientations of the figurine are then reconstructed inside the virtual Stone Age cabin, with similar elements.

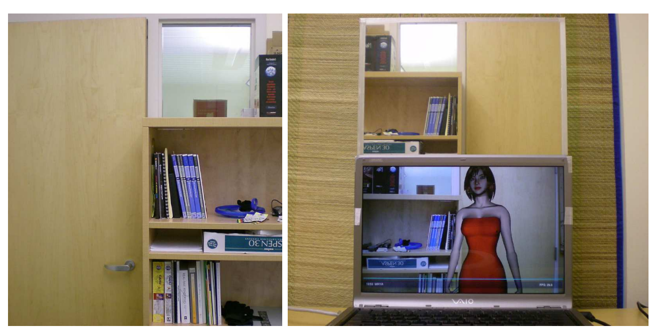

Figure 8: A virtual mirror: adding a virtual character to a real office scene (left). A real mirror is mounted on top of a laptop, providing a reference view. The laptop operates in a virtual mirror mode, showing the same scene, augmented with an animated 3D character.

Special Edition on Artificial Intelligence in Mobile Systems, 25(5), July 2004.

[12] G. Kim. Loosely Coupled Mixed Reality. *Proceedings of Mixed Reality Entertainment and Art Workshop*, ISMAR'07, Japan, Nara, 2007.

[13] Y. Kitamura, R. Rong, Y. Hirano, K. Asai, F. Kishino. Video agent: interactive autonomous agents generated from real-world creatures. *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology (Bordeaux, France, October 27 - 29, 2008). VRST '08*. ACM, New York, NY, pp. 30-38, 2008.

[14] T. Kumagai, J. Yamashita, O. Morikawa, K. Yokoyama, S. Fujimaki, T. Konishi, H. Ishimasa, H. Murata, K. Tomoda. Distance education system for teaching manual skills in endoscopic paranasal sinus surgery using "HyperMirror" telecommunication interface, *Proceedings of IEEE VR Conference*, Reno, Nevada 2008.

[15] 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.

[16] M. Lee, M. Lee, G. Kim. Loosely-coupled vs. tightly-coupled mixed reality: using the environment metaphorically. *Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry (VRCAI '09)*, Yokohama, Japan, pp. 345-349, 2009.

[17] B. MacIntyre, M. Lohse, J. Bolter and E. Moreno. Integrating 2D Video Actors into a 3D AR System. *Presence, Special Issue on Mixed Reality*, 11(2). pp. 189-202, April, 2002.

[18] Y. Medynskiy, S. Gov, A. Mazalek, D. Minnen. Wearable RFID for Play. *Proceedings of Tangible Play workshop, Intelligent User Interfaces conference*, January 28, 2007, Honolulu, Hawaii, USA.

[19] L.-P. Morency, A. Rahimi, T. Darrell. Adaptive View-based Appearance Model. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*. 1:803–810, 2003.

[20] 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.

[21] Y. Motokawa, and H. Saito. Support system for guitar playing using augmented reality display. *Proceedings of the 5th IEEE and ACM international Symposium on Mixed and Augmented Reality (ISMAR'06)*, Washington, DC, October 22 - 25, pp. 243-244, 2006.

[22] T. Ni, G. Schmidt, O. Staadt, M. Livingston, R. Ball, R. May. A Survey of Large High-Resolution Display Technologies, Techniques, and Applications. *Proceedings of the IEEE conference on Virtual Reality (VR'06)*, IEEE Computer Society, Washington, DC, March 25 - 29, 2006.

[23] T. Ohta. Dynamically reconfigurable multi-display environment for CG contents. *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, ACE'08*, Yokohama, Japan, p.416, 2008.

[24] 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.

[25] 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.

[26] 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.

[27] A. Sherstyuk, K. Chu, S. Joseph. Virtual Mirror for Augmented Presence Applications. *Proceedings of the 2008 international Conference on Advances in Computer Entertainment Technology (ACM ACE'08)*, p.410, Yokohama, Japan, December 3-5, 2008.

[28] A. Sherstyuk, K. Chu, S. Joseph. Virtual Roommates in Ambient Telepresence Applications. *Proceedings of the 2008 International Conference on Artificial Reality and Telexistence (ICAT'08)*, Yokohama, Japan, December 1-3, 2008.

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

[30] B. Spanlang, J.-M. Normand, M. Slater. A First Person Avatar System with Haptic Feedback. *Proceedings of the 2010 ACM Symposium on Virtual Reality Software and Technology (Hong-Kong, November 21 - 23, 2010). VRST '10*. ACM, pp. 47-50, 2010.

[31] *TelePresence Conference system*, by Cisco Systems, Inc. http://www.cisco.com.

[32] R. Tenmoku, M. Kanbara, N. Yokoya. A Wearable Augmented Reality System for Navigation Using Positioning Infrastructures and a Pedometer. *Proceedings of IEEE / ACM International Symposium on Mixed and Augmented Reality*, p. 344, 2003.

[33] D. Yu, J. Jin, S. Luo, W. Lai, Q. Huang. A Useful Visualization Technique: A Literature Review for Augmented Reality and its Application, Limitation & Future Direction. *Visual Information Communication*, L. Huang et al. (Eds.), Springer, ISBN: 978-1-4419-0311-2, 2010.