

Toward a Compelling Sensation of Telepresence: Demonstrating a portal to a distant (static) office

Wei-Chao Chen

Herman Towles

Lars Nyland

Greg Welch

Henry Fuchs

Department of Computer Science
The University of North Carolina at Chapel Hill
{ciao,towles,nyland,welch,fuchs}@cs.unc.edu



Figure 1: *Office of the Future* concept drawing.



Figure 2: Portal to a distant office. Shown here in mono, the remote office scene is normally viewed in stereo by the head-tracked user.

Abstract

In 1998 we introduced the idea for a project we call the *Office of the Future*. Our long-term vision is to provide a better every-day working environment, with high-fidelity scene reconstruction for life-sized 3D tele-collaboration. In particular, we want a true sense of presence with our remote collaborator and their real surroundings. The challenges related to this vision are enormous and involve many technical tradeoffs. This is true in particular for scene reconstruction. Researchers have been striving to achieve real-time approaches, and while they have made respectable progress, the limitations of conventional technologies relegate them to relatively low resolution in a restricted volume.

In this paper we present a significant step toward our ultimate goal, via a slightly different path. In lieu of low-fidelity dynamic scene modeling we present an exceedingly high fidelity reconstruction of a real but *static* office. By assembling the best of available hardware and software technologies in static scene acquisition, modeling algorithms, rendering, tracking and stereo projective display, we are able to demonstrate a portal to a real office, occupied today by a mannequin, and in the future by a real remote collaborator. We now have both a compelling sense of just how good it could be, and a framework into which we will later incorporate *dynamic* scene modeling, as we continue to head toward our ultimate goal of 3D collaborative, telepresence.

Keywords: Telepresence, Tele-Immersion, Virtual Reality, Collaborative Visualization, Immersive Display, Augmented Reality, Human-Computer Interface

1 Introduction

Despite several decades of unprecedented advances in computing technology, there has been little change in the number of pixels we see, or in the way we interact with the computer. Our dream, depicted in Figure 1, is that some day your office would be equipped with ceiling-mounted digital cameras and “smart” projectors that work together to support high-resolution projected imagery, human-computer interaction, and dynamic image-based modeling throughout the office [22]. We foresee a future when we will be able to interact with our colleagues in any locale just as if they were across the table in our office.

However we are still years away from the full realization of our long-term vision. Network transport issues aside, a primary challenge in realizing our vision is real-time, high quality scene reconstruction (acquisition/display of geometry and textures) of real environments. While system design is a multi-dimensional optimization task, Figure 3 illustrates the fundamental design space tradeoff between scene quality and acquisition performance. To date, researchers have largely explored the lower left quadrant of this space. One would naturally expect this activity to expand along the diagonal as the research community pushes toward the ultimate goal of high quality and real-time scene reconstruction.

We have chosen to take a slight “detour” from the normal diagonal development thrust of Figure 3. In order to capture a glimpse today of what the future may hold, we chose to remove the real-time capture constraint and to focus our efforts on recreating the most compelling visual experience possible. Even though the realization of a natural and convincing immersive environment is still quite complex, we believe that we have reached a milestone in our research by demonstrating for the first time a truly compelling *portal to a distant office* (Figure 2).

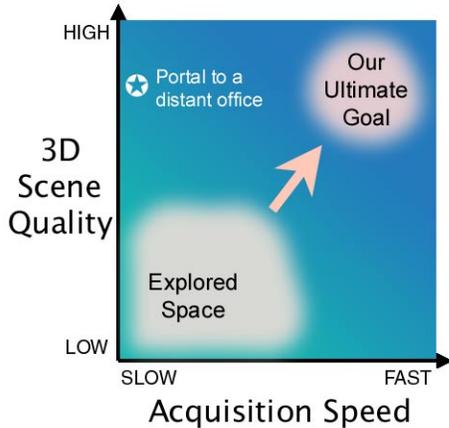


Figure 3: An important tele-immersion design space.

2 Related Work

For many years researchers have been using two-dimensional video textures and projector displays to realize immersive telepresence. Two relevant examples include [13] and [11]. They use carefully constructed geometric models for the office environment, and video-based human avatars obtained by separating the remote participants from the original background (via delta-keying). The results are presented to a head-tracked user. The TelePort display is built into a room that is carefully designed to match the rendered room [11]. The goal is for the virtual room to seem as an extension of the real room. In both instances, because the background is synthetically modeled, it is not a faithful representation of the remote collaborator’s real office. In [27], a perspectively-correct 2D live video conferencing system is proposed. This system offers high-resolution projected imagery, but because the user is not tracked the images appear geometrically distorted when the viewer is not in the system’s “sweet spot”. Billingham et al. have developed a very unusual and compelling Augmented and Virtual Reality tele-collaborative system that attaches miniature video avatars of remote collaborators to small cards in the local viewer’s real space [4]. While this may be effective for some applications, we are interested primarily in 3D reconstructions that are scaled and located to give the appearance of sitting across the table.

Three-dimensional computer vision-based techniques typically seek to use multi-baseline correspondence to estimate a model of the scene [20, 3, 9]. However, given the limitations of today’s technology, these methods generate only low-resolution depth maps. Visual hull techniques [16] can deliver approximate geometry in real time, but are primarily suited for outside-looking-in camera configurations. There is no clear extension to an inside-looking-out configuration, which would be necessary for reconstructing an environment such as an office.

The primary benefit of using some form of *image-based rendering* [14, 12] is that it is targeted at the reconstruction of real objects or environments, and does not require an explicit 3D representation of the scene. However, given the display resolution and working volume we desire, the image-based rendering cost would be quite significant compared to the use of conventional textured polygons.

Finally, one might imagine mounting a remote stereo camera rig on a head-slaved robotic arm. However, electro-mechanical limitations introduce significant latency which would result in geometric mis-registration between the local and remote scenes.

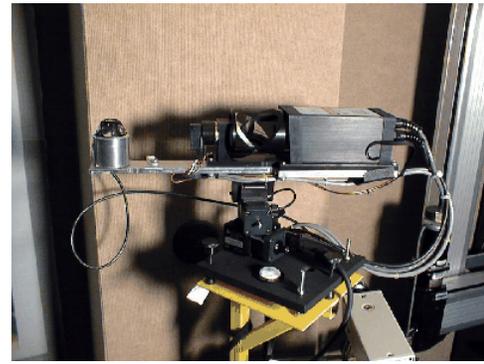


Figure 4: Laser scanner used for scene acquisition.

3 Implementation

In analyzing the task to demonstrate a portal to another office, we identified five key components of technology needed:

- **Acquisition** of a remote scene with color and depth information.
- **Modeling** and conversion of the acquired data into textured 3D geometry.
- **Tracking** of the local user’s eye positions.
- **Rendering** of the models based on the tracked user position.
- **Presentation** in stereo for the user.

In considering these needs, we further established a number of implementation priorities. First among them was the remote office environment. It had to be a real office and not a mixture of real and synthetic components. It should be acquired and presented with the best visual fidelity and resolution possible. In fact, we wanted the scene definition to exceed the display resolution so as the display resolution improved we could repeat the experiment with the same acquisition data and/or techniques. Of course, as previously discussed we knew we wanted to explore the operating space that non-realtime acquisition would afford.

In addition, our previous virtual reality work taught us the importance of accurate and low latency tracking in creating the correct visual cues without lag or swim effects. As for rendering, we wanted to achieve 30fps, and we knew we would apply the projection technology we had been working with in our *Office of the Future* research. We also hoped to apply the dual-projector, passive stereo techniques we had recently developed.

The following sections individually address the five technology areas and detail our implementation decisions and methods.

3.1 Acquisition

For scene acquisition, we selected a 3D scanning laser rangefinder system [17, 1] because of the unmatched quality and density of data provided compared to other camera-only, vision-based techniques [20, 3, 9].

This system (Figure 4), which was developed in-house, captures range data in spherical coordinates - scanning the room in elevation passes before stepping to the next azimuth position. Range data accuracy is approximately 4mm RMS at 1 milliradian angular resolution in azimuth and elevation. Scene color is acquired by replacing the scanner with a calibrated high-resolution digital camera matched to the same center of projection. Color images are taken

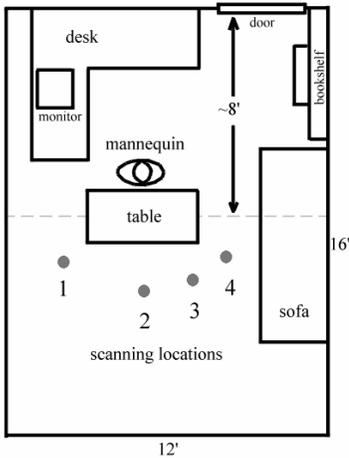


Figure 5: The layout of the office and four scanning positions.



Figure 6: The real office before scanning, with a *cardboard portal mockup* hung from the ceiling.

at azimuthal steps of 24 degrees. Acquisition time for a scene scan of 180 degrees azimuth is approximately 30 minutes.

Figure 5 illustrates the layout of the office to be scanned. We positioned a table in the middle of the room with a mannequin seated behind it. The table effectively divided the 16 ft. x 12 ft. office into two parts - an 8 ft. x 12 ft. area to be scanned and later rendered (upper portion of figure), and the area from which the laser rangefinder system scanned the scene (lower portion of figure).

To produce a complete scene definition of everything behind the table, we took four scans from different positions to minimize the possibility of missing scene data due to occlusions. The scanning positions were chosen to achieve full scene definition, making use of the knowledge that the environment would be viewed from the eye height of a seated adult, through a 4 ft. x 3 ft. portal, from a distance of approximately three feet opposite the table and mannequin. To further aid in determining the scan positions we created a temporary *cardboard portal mockup* by cutting out an appropriately sized opening in a large piece of cardboard. We hung the cardboard portal mockup from the ceiling over the table (in front of the mannequin) so that we could clearly ascertain what would be visible through the virtual portal when we projected the reconstructed scene onto a display of the same dimensions. Figure 6 shows a snapshot of the office scene with the cardboard portal mockup hung from the ceiling.

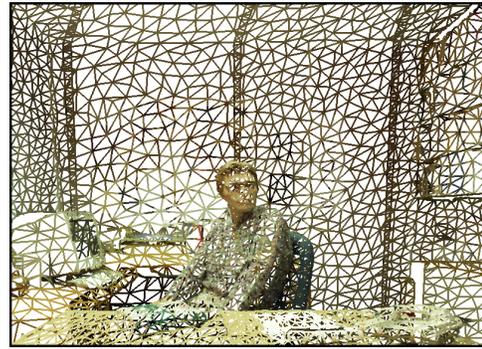


Figure 7: A portion of the polygon mesh after simplification.

Each of the four scans produced a 2924 x 1754 resolution range image with color images of similar angular resolution that covered 150 degrees in azimuth and 90 degrees in elevation. The color and range images were registered by interactively selecting correspondences between the range image and each of the color images. Because the intrinsic parameters of the color camera were already known from a priori calibration, after obtaining the extrinsic parameters with the correspondences, we were able to properly map depth values onto the color images. Final depth images were then generated by sampling depth values for each pixel in the color images.

3.2 Modeling

Several methods have been proposed for model reconstruction from range images [23, 6]. We adapted methods to combine the data from multiple range scans and high-resolution images to construct a unified polygonal model of the remote office with minimum occlusions.

After collecting the four scans of the office-mannequin scene, we employed reconstruction and simplification steps to reduce the data from more than 20 million points to a model of approximately 30K triangles (Figure 7). High-resolution color images from the scanned scene were then textured on the polygonal mesh. The images required approximately 48 MBytes of texture memory in the rendering system.

Details of the geometric simplification methodology follow.

Reconstruction Since each physical object in the office can be seen from more than one scan locations, duplicate samples can be removed for faster rendering. We choose to keep samples with better scan-quality, and discard the ones with worse quality. Samples with local surface normals closer to the gazing direction of the scanner are considered to have higher quality [17]. As the registration among different scans is not exact, some of the redundant triangles are not removed successfully. To improve rendering speed, we manually intervened to eliminate many of these duplications.

After redundant samples are removed, an input mesh is generated for each depth image using regular meshing techniques [15]. These input meshes need to be processed before simplification. First, physically disjoint objects can be falsely connected on silhouette edges. These *false triangles* can be removed by detecting empty space using scans from multiple locations [17, 6]. Although no triangles are removed incorrectly, not all false triangles can be removed successfully if the number of acquisition scans is insufficient. Therefore, we also removed triangles based on several heuristics including triangle orientation (with respect to the corresponding laser scan beam) and triangle aspect ratio [19].

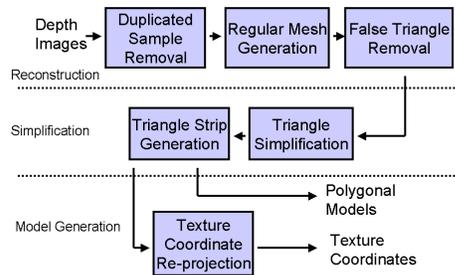


Figure 8: Steps of geometric modeling from depth images.

Simplification After false triangles and redundant triangles are removed, the resulting mesh is simplified. We have adapted a fast and robust quadric simplification method [10] to generate the smaller dataset. The density of triangles closer to the user is made higher by assigning higher quadric weights during simplification. In the end, the degree of mesh simplification is driven by the desire to maintain a high rendering frame rate. To further increase the rendering speed, the triangles are processed into triangle strips [8].

Figure 8 illustrates the steps for constructing the geometric model from the depth images. Since the complexity of our model is driven by the rendering performance, the modeling process is very lengthy and currently requires much user intervention. As stated, the majority of time was spent on manually removing duplicated samples that were not automatically deleted due to registration errors among the different scanned views.

3.3 Tracking

Our application is more sensitive to tracking than typical VR applications, as the rendered scene is observed together with the real environment. In this situation, tracker latency and error can translate into dynamic and static registration error, respectively [2]. Dynamic registration errors cause the moving user to perceive shearing in the rendered scene, and these factors destroy the illusion of presence.

We chose to use the UNC-developed HiBall™ Wide-Area Tracking System [25, 26, 1] because of its superior performance characteristics compared to other wide-area trackers such as magnetic based systems. The HiBall tracker provides new estimates of tracker position and orientation over 1500 times per second with a measured RMS position and orientation noise of less than 0.5 millimeters and 0.02 degrees, respectively. The system uses a array of infrared LED beacons mounted in the overhead ceiling panels and an optical sensor (the HiBall) mounted to an adjustable headband that the user wears (Figure 9).

3.4 Rendering

The triangles that have been processed into triangle strips are rendered using *projective textures* to address texture distortion. Since our scene is largely diffuse, in order to increase the rendering speed, we currently do not exploit view-dependent texture-mapping techniques such as [7]. The simplified model is currently rendered in stereo on two pipes of SGI InfiniteReality2, one for each eye. On two genlocked graphics pipes, the current frame rate is 30fps for each eye.

Since we are using roughly-aligned projector pairs, the displayed images of the two projectors are not aligned and need to be calibrated. By finding at least four corresponding pixels between the

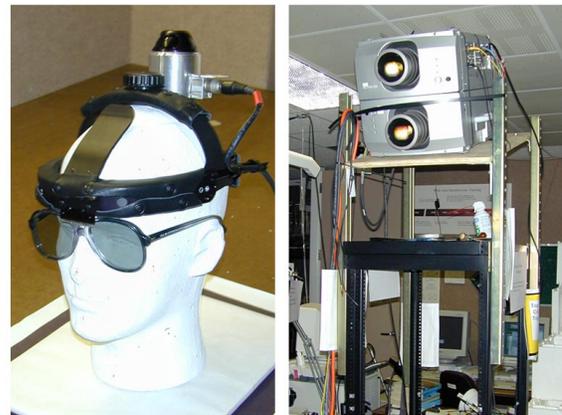


Figure 9: Left: Head-mounted optical tracker and separate passively-polarized stereo glasses. Right: Circularly-polarized stereo projector pair.

two projectors, we can evaluate the 3D collineation/homography matrix [21] that relates the two images. In practice, we find the image coordinates for each projector that map to the same four points on the corners of the display screen and solve for two collineation matrices. These matrices, $H_{left-eye}$ and $H_{right-eye}$, are then used to modify the viewing transformation of their respective view transform so as to achieve projector pair alignment to within a sub-pixel error threshold. This step introduces no extra rendering cost.

Coordinate systems need to be registered among modeling, tracking and physical world spaces. The transformation from the tracker to the world space is solved by finding the tracker readings on the corners of the physical screen. Scene modeling space is registered to the world coordinate with the transformation by specifying the location and orientation of the physical window inside the model. While orientation and positioning of the virtual office is critical, equally important is accurate scale. Inter-pupillary distance (IPD) and eye positions are critical in resolving the scale of the scene. We currently measure the user's IPD before using our system. Combined with the knowledge of the tracker's orientation and position on the headband, we can estimate the position of each eye relative to the tracker.

Given the above transformation, at each screen update we read the most recent tracker position/orientation from the networked HiBall™ tracker server and calculate the virtual eye position inside the model coordinate system. A view frustum for each eye is formed by connecting each virtual eye position to the four corners of the screen and making the projection plane parallel to the virtual location of the screen. Finally, since the collineation matrices effectively map the corners of the computed images to the pixel coordinate corresponding to the physical corners of the screens, the two matrices $H_{left-eye}$ and $H_{right-eye}$ are applied (respectively) to the left-eye and right-eye transformation matrix stacks.

To verify the registration accuracy of our system, we rendered virtual spheres relative to know positions of a set of hanging plumb weights as shown in Figure 10. As a result of remaining rendering latency, some small dynamic mis-registration is observable for a fast-moving user, but the static registration is quite accurate.

3.5 Presentation

To convincingly present the virtual office scene, it must be displayed in life-size, stereo, and properly registered with the real world. We chose to render the stereo image with two LCD pro-



Figure 10: Registration verification between real physical objects (plumb weights) and computer-generated spheres below them.

jectors, one for each eye, fitted with circularly polarizing filters. This approach simultaneously presents left and right eye views at the maximum rate, rather than in a time-division multiplexed manner [5] or by sharing the spatial resolution [24] of a single projector. In our solution, the user wears only a lightweight pair of passively polarized glasses (Figure 9). Though all LCD projectors are linearly polarized, not all projectors are built with the RGB imagers polarized in the same linear plane. By selecting a projector that does, we were simply able to add a $1/4\lambda$ retarder filter in front of the lens to produce circularly polarized light. By orienting the filter on the respective projectors 90 degrees apart, the polarization can be made clockwise for one eye and counter-clockwise for the other - thus creating the needed stereo separation. Stereo with circular polarization, rather than simple linear polarization, provides the advantage that left-right eye discrimination is independent of the head angle of the user and passive glasses. We used a pair of Sharp XG-3000U projectors augmented with polarization filters we assembled, and polarized glasses from StereoGraphics Corporation. To preserve polarization, the display surface is covered with a silver lenticular screen fabric (GN92448) from Draper, Inc.

Our rendering process does not account for nonlinear distortions (such as radial distortions) in the projected imagery. However we are able to adjust the projectors to minimize such distortions before calibration. In particular, we observed that while adjusting the zoom for the projectors, the type of radial distortion changed from barrel-type to pincushion-type, with a "sweet spot" in between where minimal distortion can be reached. Such an approach in practice alleviated the problem significantly.

Other aspects of the viewing/presentation environment that called for some judicious decisions were related to the size and position of the *projected window*. The size of 4 ft. \times 3 ft. was selected to enable life-size display of the mannequin and provide a reasonable viewing window while maintaining high display pixel density and image brightness. The desk used under the window in the real office is the exact same desk that was scanned with the mannequin. This further ensured an additional degree of continuity from the real to the virtual worlds.

4 Initial Reactions and Lessons Learned

We were so encouraged by the initial reactions of our departmental colleagues that we immediately wanted to get a more unbiased reaction from unsuspecting outsiders. So we invited two members of the university community, both professionally experienced with conventional video teleconferencing, to evaluate our system. To the



Figure 11: Working system with the tracker placed on the user's head and rendering in stereo mode.

question of how this system compared with conventional tele-video, the first subject responded,

"It's less clear in some respects, but of course, one of the things that's different is I can lean into [it, and] look into the picture where in teleconferencing you can't."

The second subject was more emotional:

"It's a bit eerie it's so realistic... it looks like somebody took a chain saw and cut a hole in the wall and he's on the other side..."

Since these initial tests and encouraging words, twenty-five or more persons have also had an opportunity to experience the demo. From these sessions, we have learned several important lessons that will be relevant when we conduct scientifically-controlled user studies of the system's effectiveness. *First*, we have seen evidence that a person's initial impressions of the system may be negatively affected by whether their first exposure is from the head-tracked, stereo position of the active participant or as an observer standing in the background only wearing stereo glasses. We speculate that if the initial experience is that of a casual observer one may become pre-biased compared to the person that first or only has a head-tracked experience. This is based on the knowledge that only the head-tracked user is seeing a perspectiveally correct view and probably more importantly a view that is responding naturally to their head-motions.

Secondly, we have found that the quality of depth perception varies significantly between individuals and is truly impacted by our estimate of interpupillary distance and geometric relationship to the head-mounted tracking system.

Lastly, when we run the stereo display with one SGI graphics pipe and the frame rate drops from 30fps to 15fps, the experience is *much less compelling*. Some of us speculate that the perceived effectiveness may fall apart rapidly when certain performance levels (e.g., frame rate) are not maintained. In fact, we are concerned that some researchers who have built similar systems in the past may not have been as encouraged simply because system performance fell below some acceptable psychological threshold. Brooks, Whitton, et al. here at UNC are developing sound scientific methods for measuring perceptual effectiveness and sensitivity to variations in a variety of system parameters.

Figure 11 and the short compressed movie on the CD-ROM distribution show the system in operation.

5 Conclusions and Future Work

When the first subject immediately started to move into the scene and the second subject spontaneously described it as cutting a hole into the wall of the next office, we realized that we may have reached our goal - that of giving a glimpse of what future, naturally immersive environments are going to be like.

The major challenge now is to advance this research from a proof-of-concept demo of a high quality, *static* scene to a fully working high-quality, *dynamic* 3D telepresence system. In that spirit, we are currently:

- Developing real-time point cloud, geometric meshing algorithms
- Investigating fast, intuitive methods for calibrating the eye positions of individual users with respect to head-mounted trackers
- Exploring tracker technology which does not encumber the user (no headware and wireless)
- Applying multiple, intensity blended, front-projector arrays for increased resolution, area and brightness
- Seeking new stereo projectors to replace our current solution that is dependent on a projector polarization characteristic that is no longer available.
- Exploring new algorithms and hardware architectures for real-time image-based rendering
- Discussing more extensive user-studies for quantifying the perceptual quality of our results and sensitivity to various system parameters

But first and foremost, we realize that the laser scanning technology we used in this experiment does not directly promise offer a real-time solution. Therefore, as part of the National Tele-Immersion Initiative we are actively incorporating real-time reconstruction techniques, developed at the University of Pennsylvania's GRASP Lab [18] under the leadership of Ruzena Bajcsy and Kostas Daniilidis, with our high-fidelity, non-realtime scanned office environments. We are also exploring the application of real-time, scene acquisition technologies based on visual hull algorithms developed by Leonard McMillan and his research team at MIT [16]. While today's real-time capabilities support only limited working volumes and resolution, these investigations are necessary steps toward our ultimate goal.

Although many years of work remain, we are increasingly optimistic that the user-computer interface will no longer be limited to a desktop monitor, but will provide a more compelling, immersive environment for an increasing variety of applications, in particular those that bring distant people closer together.

6 Acknowledgments

Support for this research is generously provided by the *National Tele-Immersion Initiative (NTII)* sponsored by Advanced Network and Services, Inc. and NSF Cooperative Agreement no. ASC-8920219: "Science and Technology Center for Computer Graphics and Scientific Visualization". We also thank the Intel Corporation for their generous equipment donations as part of *Technology for Education 2000*.

We would also acknowledge the support of our fellow research team members at UNC-CH, listed at www.cs.unc.edu/Research/stc/office/index.html, and graciously credit and thank David McAllister and Steven Matuszek for their time and software utilities used in processing the rangefinder data. We thank Andrew Nashel for his help in putting together the video for the paper. We thank John Thomas and Kurtis Keller for their engineering assistance, and finally we thank Andrei State for his visionary sketch (Figure 1).

References

- [1] 3rdTech, Inc. <http://www.3rdtech.com/>.
- [2] Ronald Azuma and Gary Bishop. Improving Static and Dynamic Registration in an Optical See-Through HMD. In *SIGGRAPH '94 Conference Proceedings*, pages 197–204, July 1994.
- [3] Ruzena Bajcsy, Reyes Enciso, Gerda Kamberova, and Lucien Nocera. 3-D Reconstruction of Environments for Virtual Collaboration. *Proc. 4th IEEE Workshop on Applications of Computer Vision*, 1998.
- [4] Mark Billinghurst. Shared Space: Collaborative Augmented Reality. In *SIGGRAPH '99 Conference abstracts and applications*, pages 178–178, 1999.
- [5] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE. In *SIGGRAPH '93 Conference Proceedings*, volume 27, pages 135–142, August 1993.
- [6] Brian Curless and Marc Levoy. A Volumetric Method for Building Complex Models from Range Images. In *SIGGRAPH '96 Conference Proceedings*, pages 303–312, August 1996.
- [7] Paul E. Debevec, Yizhou Yu, and George D. Borshukov. Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping. In *Eurographics Rendering Workshop*, pages 105–116, 1998.
- [8] Francine Evans, Steven S. Skiena, and Amitabh Varshney. Optimizing Triangle Strips for Fast Rendering. In *IEEE Visualization '96*, October 1996.
- [9] O. Faugeras. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, Massachusetts, 1993.
- [10] Michael Garland and Paul S. Heckbert. Surface Simplification Using Quadric Error Metrics. In *SIGGRAPH '97 Conference Proceedings*, pages 209–216, August 1997.
- [11] Simon Gibbs, Constantin Arapis, and Christian J. Breiteneder. TELEPORT-Towards Immersive Copresence. *Multimedia Systems*, 1999.
- [12] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The Lumigraph. In *SIGGRAPH '96 Conference Proceedings*, pages 43–54, August 1996.
- [13] Vali Lalioti, Christophe Garcia, and Frank Hasenbrink. Virtual Meeting in Cyberstage. In *Proceedings of the ACM Symposium on Virtual reality software and technology*, pages 205 – 212, 1998.
- [14] Marc Levoy and Pat Hanrahan. Light Field Rendering. In *SIGGRAPH '96 Conference Proceedings*, pages 31–42, August 1996.
- [15] William R. Mark, Leonard McMillan, and Gary Bishop. Post-Rendering 3D Warping. In *1997 Symposium on Interactive 3D Graphics*, pages 7–16, April 1997.
- [16] Wojciech Matusik, Christopher Buehler, Ramesh Raskar, Steven Gortler, and Leonard McMillan. Image-based Visual Hulls. In *SIGGRAPH 2000 Conference Proceedings*, July 2000.

- [17] David K. McAllister, Lars F. Nyland, Voicu Popescu, Anselmo Lastra, and Chris McCue. Real-Time Rendering of Real-World Environments. *Eurographics Rendering Workshop 1999*, June 1999.
- [18] J. Mulligan and K. Daniilidis. View-independent Scene Acquisition for Tele-Presence. In *Proceedings of International Symposium on Augmented Reality*, October 2000.
- [19] Kari Pulli, Michael Cohen, Tom Duchamp, Hugues Hoppe, Linda Shapiro, and Werner Stuetzle. View-based Rendering: Visualizing Real Objects from Scanned Range and Color Data. In *Eurographics Rendering Workshop 1997*, pages 23–34, June 1997.
- [20] Peter W. Rander, P.J. Narayanan, and Takeo Kanade. Virtualized Reality: Constructing Time-Varying Virtual Worlds from Real World Events. *IEEE Visualization '97*, pages 277–284, November 1997.
- [21] Ramesh Raskar. Immersive Planar Display using Roughly Aligned Projectors. In *IEEE Virtual Reality 2000*, March 2000.
- [22] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays. In *SIGGRAPH '98 Conference Proceedings*, pages 179–188, 1998.
- [23] Greg Turk and Marc Levoy. Zippered Polygon Meshes from Range Images. In *SIGGRAPH '94 Conference Proceedings*, pages 311–318, July 1994.
- [24] VRex, Inc. <http://www.vrex.com/>.
- [25] Greg Welch and Gary Bishop. SCAAT: Incremental Tracking with Incomplete Information. In *SIGGRAPH '97 Conference Proceedings*, pages 333–344, August 1997.
- [26] Greg Welch, Gary Bishop, Leandra Vicci, Stephen Brumback, and Kurtis Keller. The HiBall Tracker: High-Performance Wide-Area Tracking for Virtual and Augmented Environments. In *Symposium on Virtual Reality Software and Technology*, December 1999.
- [27] Ruigang Yang, Michael S. Brown, W. Brent Seales, and Henry Fuchs. Geometrically Correct Imagery for Teleconferencing. In *ACM Multimedia '99 Conference Proceedings*, pages 179–186, 1999.