Beyond the Desktop Metaphor: Toward More Effective Display, Interaction, and Telecollaboration in the Office of the Future via a Multitude of Sensors and Displays

Henry Fuchs

Department of Computer Science
University of North Carolina at Chapel Hill, USA

Abstract.

We are engaged in a long-term project to improve personal productivity for computer-related activities and tele-collaboration in an office environment of the future. Personal computer-related activities, we believe will be enhanced by capability to project imagery on any surface in the office, that together with precise head and eye-tracking, will enable head-tracked stereo imagery to be added to the user's views of his/her office environment --- creating a 3D immersive generalization of the now ubiquitous 2D desktop metaphor as the principal human-computer interface. We plan to realize this kind of system by mounting many video projectors and video cameras around the room, especially around the ceiling. The projectors may provide the only source of light in the room and will allow detailed imagery to be projected (almost) everywhere in the office. In order to generate the appropriate imagery, however, a detailed 3D map of the changing office environment needs to be acquired. This will be acquired by measuring, with synchronized cameras and projectors, the precise 3D location(s) of the surface(s) light up by each pixel of each projector. Local collaboration will be enhanced by tracking each of several individuals in the office and generating (by time-division multiplexing or by other means), a stereo image pair appropriate for each individual. Objects under design may be displayed, for each individual, from his/her own perspective and to his/her own specifications of interest. TELE-collaboration activities, we believe will be enhanced by having such an enhanced office environment for each of the small group of distant collaborators, and displaying for each participant, in addition to the shared objects under discussion, some combination of the remote scenes that include the changing 3D images of each of the participants and 3D images of physical objects of joint interest. To realize many of these capabilities, each user may need to wear polarized eyeglasses to perceive proper stereo imagery. Although initial results are encouraging, numerous difficult problems remain -- how, for example, can imagery be projected onto a dark-colored surface in the room. The cost of such systems, with many projectors and cameras, image generators and image acquisition devices, may initially be prohibitively expensive, but is expected to decrease as the cost of such off-the-shelf equipment naturally decreases with increased market size. The positive psychological effects of working and interacting in such an immersive

S. Nishio, F. Kishino (Eds.): AMCP'98, LNCS 1554, pp. 30-43, 1999.

environment within a "standard" office will be so compelling, we believe, that users will not readily wish to return to working within the constraints of a 21" monitor. Much of this work is being carried out as part of a collaboration among the five sites of the NSF Science and Technology Center in Computer Graphics and Scientific Visualization (Brown, Caltech, Cornell, UNC, and Utah) and is also being carried on in collaboration with the GRASP Lab at University of Pennsylvania, and is part of the National Teleimmersion Initiative sponsored by Advanced Network and Services.

Introduction

This project originated from two influences in our department: 1) many years of research into virtual reality technologies and applications, and 2) a long tradition of working in collaboration with others near and far (rather than working alone). Long-distance collaboration in particular has been at the core of the Science and Technology Center for Computer Graphics and Visualization ("the Graphics STC"), of which UNC is one of the five sites. The five sites are widely separated within the USA, so the distant participants have no choice but to rely on whatever tools and techniques are available to enable the long-distance collaboration.

Several of the virtual reality applications have in recent years involved "augmented reality" -- visualization and interaction not just with a completely synthetic, computer-generated environment, but visualization and interaction with an environment composed of *a fusion of* the computer-generated one and the user's own physical environment (the user's environment *augmented* with a synthetic one). In several medical applications, for example, the user (a physician) sees the patient on the examining table AND this patient's image data properly registered inside. In this way the physician could, for example, identify the precise location of tumors during needle biopsy. About five years ago (1993), this vision of augmented reality was expanded to include augmentation of the user's local environment by (3D) images of distant environments, as in the linking of teachers with students in separate classrooms, and establishing remote surgical consultations between medical experts based at certain sites and physicians handling cases (such as trauma) in others.

Support for collaboration within the Graphics STC began with standard tele-conferencing technology. Later developments included an experimental shared virtual environment with "live" video images of participants' faces on crude rectangular avatars in the shared virtual environment. From these experiments, some of us concluded that remote participants would benefit not just from "live" video of the faces of the other participants, but also from 3D reconstructions of each participant and his or her nearby environment.

Recently these two interests have come together in a project supported by the National Tele-immersion Initiative funded by Advanced Network and Services. Their support, along with that of NSF and DARPA for related research, have enabled us to begin building a prototype office in the UNC graphics and image lab (built along the dimensions of a standard 12 foot x 16 foot university office) to try out these ideas of remote tele-collaboration. The near term objective is to link two rooms on the same floor for telecollaboration experiments — a small conference room and the prototype office (We are also working on linking the conference room to a distant lab.)

Design Considerations for the Office of the Future

Given that the future of tele-collaboration technology will be of most use to individual researchers working alone or with local colleagues rather than by teams working at great distances, whatever tele-collaborative system that is built is to be optimized for solo and local work. Additionally, in a world where increasingly individuals are basing their work environment at home, *every* interaction will be remote.

For display technology, we would like to use head-mounted displays, but these are not well enough developed at the present time to give a comfortable, stable image with sufficient image quality. Thus we are "forced" to consider other display technologies, ones with displays *not* mounted on the head. Among these remaining possibilities are projection-based displays: front- and rear-projection. Since rear-projection arrangements take up impractically large amounts of space for an office environment, we have been focusing on front-projection.

Next, we have had to decide *where* to project. Our answer was, onto well-specified surfaces such as walls or onto more general surfaces. At. UNC we are beginning with display onto walls, but we are eager to project images anywhere in the environment. In Raskar, *Siggraph98*, we show a method for pre-distorting the projected image in order to cancel out distortion effects due to the geometry of the surface. This approach fits in well with concepts developed here since 1993, such as extracting 3D scene descriptions from a "sea of cameras," so that in theory every available surface location can be extracted. This, in turn, has led us to an interest in real-time depth extraction techniques that take advantage of lighting control provided by the projectors. In turn, this has led us to a consideration of techniques for dynamically structured light that are not too distracting to the participants. Other considerations are support for display in stereo mode, and support for multiple participants.

Implementation and Problems

We are eager to try various parts of such a system as early as possible — in order to benefit from them and to learn the problems and limitations. Two such problems are listed below:

1) <u>shadowing</u> created by the user's body getting in the way of the projector paths. This results in parts of the image not reaching the intended surface. Although this getting-in-the-way is necessary in order to do the 3D scanning of the user's own body

surface, we think it may be very irritating to have NO displayed image on the some surface location of interest, such as the desk surface directly under a user's hand. We hope to lessen this difficulty by the careful placement of projectors that aim onto the surfaces and by allocating multiple projectors to project onto regions of particular importance and particularly susceptible to this shadowing problem.

2) <u>interreflections</u>—significant image degredation may occur in the Office from (projected) light bouncing from one surface to another, reducing contrast and readability. This effect is especially noticeable near inside lit corners of the Office. Since we intend to control all the light in the room through projectors, we can predict some of these scattering effects, but we can't control them

We have begun to use these concepts in simplest form, using projectors mounted on the ceiling and projecting onto walls (actually, wall-mounted white boards). Professor Gary Bishop (UNC) has already converted his private office this way (see his UNC web page photos at http://www.cs.unc.edu/~gb/office.htm) and uses it daily. In the future, for telecollaboration, we envision a system consisting of only a few sites (4-5, each with only 1 or 2 people) sending images to one another. Each collaborator at work would have the impression of two rooms from different sites being adjacent to a picture window in between. With more than two sites involved, this impression would be expanded, in that the offices would appear to be in grid arrangement and meeting at the corners. Each site, therefore, can see the others with the impression of intimacy, with their tables and chairs pushed up to the viewing surfaces.

Extractions of 3D suirface points are done by imperceptible structured light and cameras. We hope to acquire surface information (location and color) of every millimeter of lighted surface in the scene and send any changes (from people moving, etc.) to each of the other sites. In order to render the proper image, the precise location of the participant's eyes must be known.

Current Status of Project

We have just begun operating eight projectors simultaneously in our 12-foot by 16-foot prototype office, and experimenting with (very near) "telecollaboration" between a single participant in that office and a group of participants in a small conference room just a few meters away. The projectors present a panorama of images from the conference room acquired by a single cluster of 10 to 12 cameras. This specialized cluster of cameras was developed in collaboration with the Alpha_1 Project and the Advanced Manufacturing Lab at the University of Utah [Chi, et al, 1998].

We have also developed proof-of-concept demonstrations for imperceptible structured light and for correction of geometric distortion induced by nonplanar geometry of the display surfaces. We have also demonstrated stereo display using micro-mirror based projectors — called Digital Light Projection (DLP) by its developer, Texas Instruments.

Our collaborators at the University of Pennsylvania have developed techniques to extract precise depth and color maps from increasingly complex room environments.

Conclusions

Although various pieces of technologies are just beginning to be demonstrable, we see no fundamental barrier to realizing the long-term plans described above. For example, the idea of outfitting a room with dozens of projectors and digital cameras appears daunting at first, but we are encouraged that both of these items (cameras or projectors) are widely available, with steadily declining costs. We foresee a time when such configurations will be widely affordable. Initial results of displaying many coordinated images simultaneously encourage us in the belief that, once an individual experiences such a rich, interactive environment, he will never again be satisfied with a 19” display screen. Once telecollaboration can be experienced in a convincing, surround fashion, today’s conventional teleconferencing with a small window on a single screen, showing a distant participant’s “talking head”, will appear as appealing as a console teletype for computer interaction after experiencing a multi-windows system.

Acknowledgement

This continuing project is the result of active collaboration among many individuals; Greg Welch coordinated much of the initial work in his capacity as the UNC site coordinator of the Graphics STC. Herman Towles has led most of the implementation effort since joining us early this year; Graduate research assistants Matt Cutts, Ramesh Raskar, Adam Lake, and Lev Stesin developed most of the initial ideas and built many of the first system components. Graduate research assistants Ruigang Yang, Aditi Majumder, David Marshburn and Wei-Chao Chen have joined the team this fall 1998 and have already contributed significantly to various camera calibration, projector calibration and audio subsystems. Visiting Professor Brent Seales and graduate research assistant Mike Brown from the University of Kentucky are considerably strengthening our computer vision, image processing and calibration capabilities.

Much of this work is being carried out as part of a collaboration among the five sites of the NSF Science and Technology Center in Computer Graphics and Scientific Visualization (Brown, Caltech, Cornell, UNC, and Utah) and is also being carried on in collaboration with the GRASP Lab at University of Pennsylvania. Recent support from Advanced Network and Services, as part of their National Tele-immersion Initiative, has enabled us to transition this year from developing proof of concept components to building a full office-sized prototype.

Bibliography

Aliaga, J., D. Cohen, A. Wilson. H. Zhang, C. Erickson, K. Hoff, T. Hudson, W. Stüerzlinger, E. Baker, R. Bastos, M. Whitton, F. Brooks, D. Manocha.1998. "A Framework for the Real-Time Walkthrough of Massive Models," UNC Computer Science Technical Report TR98-013, University of North Carolina at Chapel Hill, March 1998.

Azuma, R., Bishop, G. Improving Static and Dynamic Registration in an Optical See-through HMD. Proceedings of SIGGRAPH 94 (Orlando, Florida, July 24-29, 1994). In *Computer Graphics* Proceedings, Annual Conference Series, 1994, ACM SIGGRAPH, pp. 197-204.

Bajaj, C.L., F. Bernardini, and G. Xu. "Automatic reconstruction of surfaces and scalar fields from 3D scans," SIGGRAPH 95 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addison-Wesley, pp. 109-118, August 1995.

Bowen, Loftin, R. "Hands Across the Atlantic," IEEE Computer Graphics and Applications, Vol. 17, No. 2, pp. 78-79, March-April 1997.

Bryson, Steve, David Zeltzer, Mark T. Bolas, Bertrand de La Chapelle, and David Bennett. "The Future of Virtual Reality: Head Mounted Displays Versus Spatially Immersive Displays," SIGGRAPH 97 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addison-Wesley, pp. 485-486, August 1997.

Buxton, W., Sellen, A. & Sheasby, M. "Interfaces for multiparty videoconferencing," In K. Finn, A. Sellen & S. Wilber (Eds.). Video Mediated Communication. Hillsdale, N.J.: Erlbaum, pp. 385-400, 1997.

Capin, Tolga K., Hansrudi Noser, Daniel Thalmann, Igor Sunday Pandzic and Nadia Magnenat Thalman. "Virtual Human Representation and Communication in VLNet," IEEE Computer Graphics and Applications, Vol. 17, No. 2, pp. 42-53, March-April 1997.

Chi, Vern, Matt Cutts, Henry Fuchs, Kurtis Keller, Greg Welch, Mark Bloomenthal, Elaine Cohen, Sam Drake, Russ Fish, Rich Riesenfeld. 1998. "A Wide Field-of-View Camera Cluster", University of North Carolina at Chapel Hill, Dept of Computer Science, Technical Report TR98-018.

Chien, C.H., Y.B. Sim, and J.K. Aggarwal. "Generation of volume/surface octree from range data," The Computer Graphics Society Conference on Computer Vision and Pattern Recognition, pp. 254-260, June 1988.

Conner, D.B, Cutts, M., Fish, R., Fuchs, H., Holden, L., Jacobs, M., Loss, B., Markosian, L., Riesenfeld, R., and Turk, G. "An Immersive Tool for Wide-Area

Collaborative Design," TeamCAD, the First Graphics Visualization, and Usability (GVU) Workshop on Collaborative Design. Atlanta, Georgia, May 12-13, 1997.

Connolly, C.I. "Cumulative generation of octree models from range data," Proceedings, Int'l. Conference Robotics, pp. 25-32, March 1984.

Cruz-Neira, Carolina, Daniel J. Sandin, and Thomas A. DeFanti. "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE," Computer Graphics, SIGGRAPH Annual Conference Proceedings, 1993.

Cutts, Matt, Henry Fuchs, Adam Lake, Ramesh Raskar, Lev Stesin, and Greg Welch. 1998. "The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays," SIGGRAPH 98 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addison Wesley, July 1998, Orlando, FL, USA

Curless, Brian, and Marc Levoy. "A Volumetric Method for Building Complex Models from Range Images," SIGGRAPH 96 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addison-Wesley. pp. 303-312, 1996.

DePiero, F.W., and Trivedi, M.M., "3-D Computer Vision using Structured Light: Design, Calibration, and Implementation Issues," Advances in Computers(43), 1996, Academic Press, pp.243-278

Dias, José Miguel Salles, Ricardo Galli, António Carlos Almeida, Carlos A. C. Belo, and José Manuel Rebordã. "mWorld: A Multiuser 3D Virtual Environment," IEEE Computer Graphics and Applications, Vol. 17, No. 2., pp. 55-64, March-April 1997.

Fuchs, Henry, Gary Bishop, Kevin Arthur, Leonard McMillan, Ruzena Bajcsy, Sang Lee, Hany Farid, and Takeo Kanade. "Virtual Space Teleconferencing Using a Sea of Cameras," Proceedings of the First International Symposium on Medical Robotics and Computer Assisted Surgery, (Pittsburgh, PA.) Sept 22-24, 1994.

Gajewska , Hania , Jay Kistler, Mark S. Manasse, and David D. Redell. "Argo: A System for Distributed Collaboration," (DEC, Multimedia '94)

Gibbs, Simon, Constantin Arapis and Christian J. Breiteneder. "TELEPORT-Towards Immersive Copresence," to appear in ACM Multimedia Systems Journal, 1998.

Hilton, A., A.J. Toddart, J. Illingworth, and T. Windeatt. "Reliable surface reconstruction from multiple range images," In Fouth European Conference on Computer Vision, Volume 1, pp. 117-126. April 1996.

Holloway, R. "Registration Errors in Augmented Reality Systems," PhD Thesis. University of North Carolina at Chapel Hill, 1995.

Holmes, Richard E. "Common Projector and Display Modules for Aircraft Simulator Visual Systems," Presented at the IMAGE V Conference, Phoenix, AZ, June 19-22, pp. 81-88, 1990.

Hornbeck, Larry J., "Deformable-Mirror Spatial Light Modulators,"Proceedings SPIE, Vol. 1150, Aug 1989.

Hornbeck, Larry J., "Digital Light Processing for High-Brightness High-Resolution Applications," Available from http://www.ti.com/dlp/docs/business/resources/white/hornbeck.pdf, 1995.

Ichikawa, Y., Okada, K., Jeong, G., Tanaka, S. and Matsushita, Y.: "MAJIC Videoconferencing System: Experiments, Evaluation and Improvement'," In Proceedings of ECSCW'95, pp. 279-292, Sept. 1995.

Ishii, Hiroshi, Minoru Kobayashi, Kazuho Arita. "Iterative Design of Seamless Collaboration Media," CACM, Volume 37, Number 8, pp. 83-97, August 1994.

Kanade, Takeo and Haruhiko Asada. "Noncontact Visual Three-Dimensional Ranging Devices," Proceedings of SPIE: 3D Machine Perception. Volume 283, Pages 48-54. April 23-24, 1981.

Kanade, Takeo, Hiroshi Kano, Shigeru Kimura, Atsushi Yoshida, Kazuo Oda. "Development of a Video-Rate Stereo Machine," Proceedings of International Robotics and Systems Conference (IROS ’95). pp. 95-100, Pittsburgh, PA., August 5-9, 1995.

Lamotte, Wim, Eddy Flerackers, Frank Van Reeth, Rae Earnshaw, Joao Mena De Matos. Visinet: Collaborative 3D Visualization and VR over ATM Networks. IEEE Computer Graphics and Applications, Vol. 17, No. 2, pp. 66-75, March-April 1997.

Lehner, Valerie D., and Thomas A. DeFanti. "Distributed Virtual Reality: Supporting Remote Collaboration in Vehicle Design," IEEE Computer Graphics and Applications, Vol. 17, No. 2, pp. 13-17, March-April 1997.

Lyon, Paul. "Edge-blending Multiple Projection Displays On A Dome Surface To Form Continuous Wide Angle Fields-of-View," pp. 203-209. Proceedings of 7th I/ITEC, 1985.

Macedonia, Michale R. and Stefan Noll. "A Transatlantic Research and Development Environment," IEEE Computer Graphics and Applications, Vol. 17, No. 2, pp. 76-82, March-April 1997.

Mandeville, J., T. Furness, M. Kawahata, D. Campbell, P. Danset, A. Dahl, J. Dauner, J. Davidson, K. Kandie, and P. Schwartz. "GreenSpace: Creating a Distributed

Virtual Environment for Global Applications," Proceedings of IEEE Networked Virtual Reality Workshop, 1995.

Milgram, P and Kishino, F. "A taxonomy of mixed reality visual displays", *IEICE (Institute of Electronics, Information and Communication Engineers) Transactions on Information and Systems*, Special issue on Networked Reality, Dec. 1994.

Milgram, P., Takemura, Utsumi and F. Kishino. "Augmented Reality: A class of displays on the reality-virtuality continuum". *SPIE Vol. 2351-34, Telemanipulator and Telepresence Technologies,* 1994.

Nayar, Shree, Masahiro Watanabe, Minori Noguchi. "Real-time Focus Range Sensor," Columbia University, CUCS-028-94.

Neumann, Ulrich and Henry Fuchs, "A Vision of Telepresence for Medical Consultation and Other Applications," Proceedings of the Sixth International Symposium on Robotics Research, Hidden Valley, PA, Oct. 1-5, 1993, pp. 565-571.

Ohya, Jun, Kitamura, Yasuichi, Takemura, Haruo, et al. "Real-time Reproduction of 3D Human Images in Virtual Space Teleconferencing," IEEE Virtual Reality International Symposium. Sep 1993.

Raskar, Ramesh, Matt Cutts, Greg Welch, Wolfgang Stüerzlinger. "Efficient Image Generation for Multiprojector and Multisurface Displays," University of North Carolina at Chapel Hill, Dept of Computer Science, Technical Report TR98-016, 1998.

Raskar, Ramesh, Henry Fuchs, Greg Welch, Adam Lake, Matt Cutts. "3D Talking Heads : Image Based Modeling at Interactive rate using Structured Light Projection,"University of North Carolina at Chapel Hill, Dept of Computer Science, Technical Report TR98-017, 1998

Segal, Mark, Carl Korobkin, Rolf van Widenfelt, Jim Foran, and Paul E. Haeberli. 1992. "Fast shadows and lighting effects using texture mapping," SIGGRAPH 92 Conference Proceedings, Annual Conference Series, ACM SIGGRAPH, Addison Wesley, volume 26, pages 249-252, July 1992.

Slater, M. and M. Usoh. (1994). "Body-Centered Interaction in Immersive Virtual Environments," in N. Magnenat Thalmann and D. Thalmann (Eds.), *Artificial Life and Virtual Reality* (pp. 125-148). London: Wiley

State, A., Hirota, G., Chen, D.T., Garrett, W.F., Livingston, M.A. "Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking". Proceedings of SIGGRAPH `96 (New Orleans, LA, August 4-9, 1996).

In *Computer Graphics* Proceedings, Annual Conference Series, 1996, ACM SIGGRAPH.

Tsai , Roger Y. "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, pp. 364-374, 1986.

Underkoffler, J. "A View From the Luminous Room," Personal Technologies, Vol. 1, No. 2, pp. 49-59, June 1997.

Underkoffler, J., and Hiroshi Ishii. "Illuminating Light: An Optical Design Tool with a Luminous-Tangible Interface," Proceedings of CHI '98, ACM, April 1998.

Fig. 1. Artist conception of the Office of the Future, illustrating a single local participant collaborating with two participants at one remote site and a single participant at one remote site collaborating on the design and manufacture of a Head Mounted Display. The geometric model of therelevant parts of a face are shown next to the boxlike Head Mounted Display so the esigners remain aware of the physical constraints on the design imposed by the contours of the human face.

Fig. 2. The Office scene during projector calibration procedure, showing the results of automatic blending of overlapped projected image regions.

Fig. 3. The Office scene during display of CAD data but without images of distant collaborators.

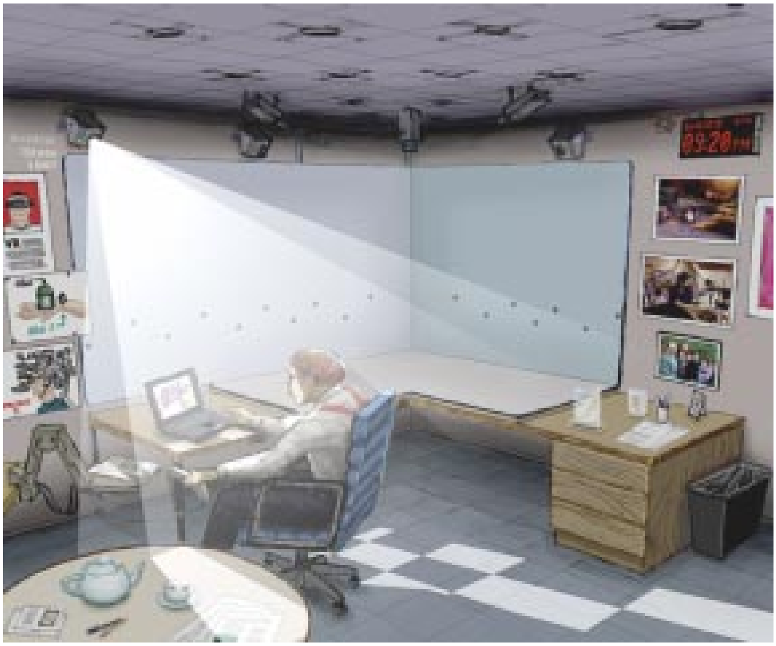

Fig. 4. The Office during 3D scene scan showing projection of structured light onto the user. We expect to use imperceptible structured light techniques so that the checkerboard pattern will not be perceptible to the user.

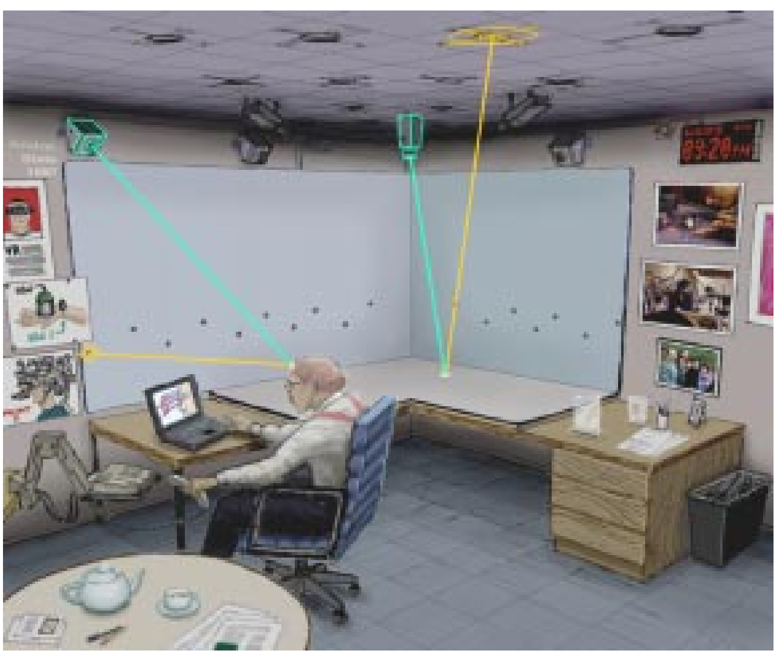

Fig. 5. The Office during extraction of points of particular interest as part of 3D scene scan, also using imperceptible light techniques so the points will not be perceptible to the user.

Fig. 6. The Office after a scene scan outlines regions of particular interest -- in this case, points which have changed since the previous night.

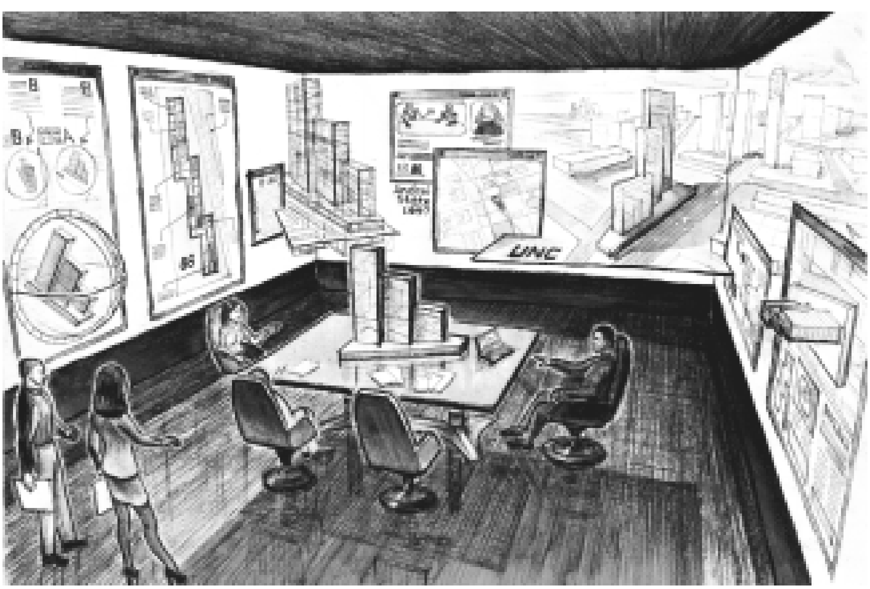

Fig. 7. Another application in which there are multiple users. If done with stereo, then 2n images for n users, a distinct stereo pair for each user's point of view.

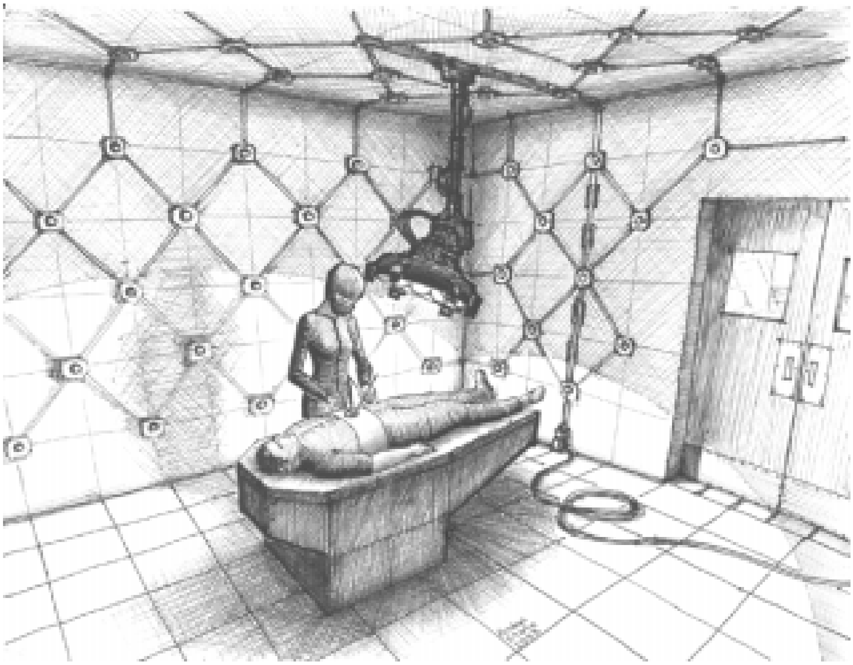

Fig. 8. A medical application with a local surgeon showing the sea of cameras around the operating room.

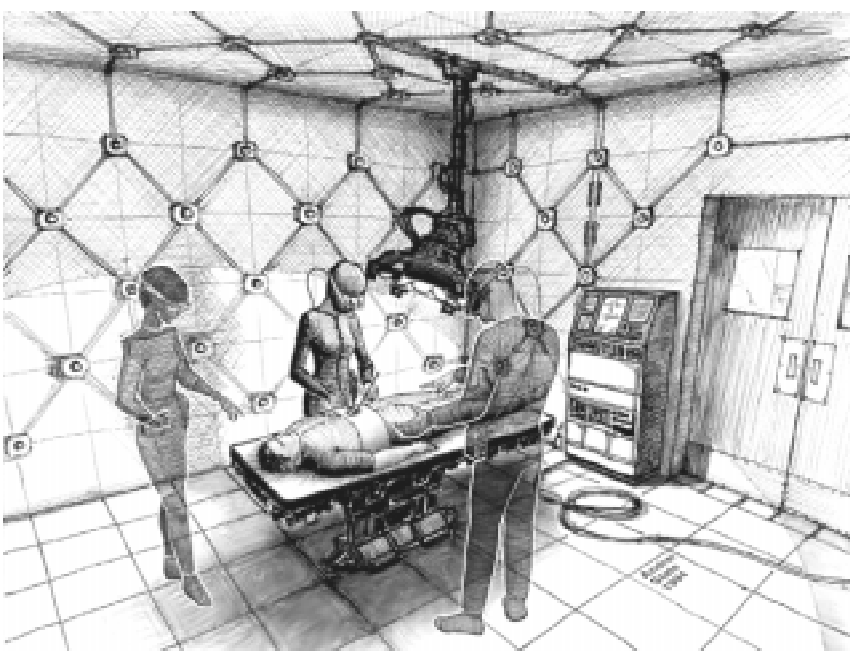

Fig. 9. The same application in which the remote consultants are viewed by Head Mounted Displays, so that each participant sees the others.