Collaborative Visualization of an Archaeological Excavation

Hrvoje Benko, Edward Ishak, and Steven Feiner
Department of Computer Science, Columbia University
New York, NY 10027
{benko,ishak,feiner}@cs.columbia.edu

EXTENDED ABSTRACT

Abstract

We are developing a collaborative system for offsite visualization of an archaeological dig site through both virtual and augmented reality. Multiple users, wearing tracked, head-worn, see-through displays, can interact with the environment using tracked, instrumented gloves, a multi-user, multi-touch, projected table surface, large wall displays and tracked hand-held displays. We take advantage of our ongoing work on 3D multimodal interaction [14,20] to allow users to combine speech with head, hand, and arm gestures to aid them in their tasks. Although the dig site can be visualized as a purely virtual environment, when users collaborate using the projected table, their see-through head-worn displays allow them to see personalized overlaid material in context with the shared, projected table surface.

1. Introduction

During the summer of 2003, our research team, consisting of archaeologists, conservators, range-scanning researchers, and visualization researchers, collected a rich set of multimedia data from an archaeological excavation. The dig site was on top of Monte Polizzo, a mountain in western Sicily, where a team of archaeologists [19] have been excavating an ancient Elymian acropolis built between the 6th and 4th centuries BC. Over a span of ten days, we collected both 2D and 3D multimedia data, including 3D point clouds, video sequences of interesting events, panoramic images, and numerous high-resolution static images of objects and the overall site, using a 3D laser range scanner, a total station surveying system, and digital video and still cameras.

Since an excavation is naturally a destructive (and often unreconstructable) process, our goal has been to capture and preserve the excavation process to allow users, ranging from interested novices to experienced archaeologists, to visualize it off-site at different points in time.

Figure 1. Two users collaborate in our system using the MERL DiamondTouch table. The virtual model of the pot is shown next to the table and the 3D terrain model is shown in the background.

We present here some preliminary results in our effort to create a collaborative 3D visualization environment to aid archaeologists in their post-excavation interpretation and analysis. As shown in Figure 1, in our system, users collaborate in an augmented reality space, wearing tracked, head-worn displays to visualize 3D terrain data and embedded multimedia. They can use a tracked glove and speech to interact multimodally with the system, as well as an interactive projected table surface to collaborate by jointly navigating, searching, and viewing data on the table's 2D user interface. We have embedded within the virtual representation of the physical site multimedia content, such as 3D panoramic images, videos and stills of the excavation process, 3D models of interesting finds, and ambient audio, as well as interpretive remarks by the archaeologists.

In the remainder of this paper, we first describe the latest techniques currently used in visualization of archaeology, and review the related work that inspired the development of our system in Section 2. Next, in Section 3, we present a short application scenario as a proof of concept, followed by a description of our current implementation in Section 4. Then, in Section 5, we discuss the merits and limitations of our system. Finally, in Section 6, we present our conclusions and discuss possibilities for future work.

2. Related Work

Since we are developing for a specific domain, we divide our review of related work into two distinct areas: applications for data visualization in archaeology and collaboration techniques in augmented reality (AR) and virtual reality (VR).

2.1. Archaeological Visualization

Archaeologists currently use various kinds of written documentation, sketches, diagrams, and photographs to document the physical state of a dig site while it is being excavated. While there are many standards or guidelines for recording the state of the dig site during excavation, their main focus is to record and archive the data, rather than visualize it.

To visualize the data, most archaeologists currently rely on geographic information systems (GIS), such as ESRI's ArcGIS [10] suite of software. Recently, INTRISIS [11] has extended some capabilities of standard GIS systems and functions as a plug-in for ArcGIS. Additionally, standard computer-aided design (CAD) systems, such as AutoCAD [3], are often used for modeling and reconstruction, and are both costly and time-consuming. While both GIS and CAD contain 3D visualization capabilities, most of those systems tend to present layered 2D maps or coarse topographical terrain maps with embedded objects, sketches, and pictures. However, additional multimedia, such as audio, video, 3D high-resolution terrain scans, and panoramas, as well as detailed object models, are currently not supported.

Several research groups have explored immersive 3D visualization for archaeology. For example, the ARCHAVE project [1] was developed for use in a CAVE [8]. It consists of a human-modeled environment embedded with virtual icons representing various types of finds and has been used to determine patterns and trends of the objects found on site. Because of display limitations, all users see the view of a single tracked user, restricting the possibilities for collaborative work. Gaitatzes et al. [12] presents various VR setups in their system, ranging from Imersadesk™-based interactive plane to a CAVE-based environment, for visualizing temples and public buildings in ancient Grece.

Based on our interviews with archaeologists, much of the typical post-excavation analysis, interpretation, report writing, and additional research they do could benefit from the ability to visually integrate both 2D and 3D data into an interactive 3D space in which 3D terrain information is combined with sketches, images, video, and other multimedia, shown being captured in Figure 2. Archaeologists envision using this space for both data interpretation after they have left the field, and for field planning and preparation for the next season. In addition, remote collaboration should be possible. For example, a pottery specialist could give her interpretation of a ceramic pot *in situ* to colleagues in a different physical location. These possibilities all suggest the potential advantages of applying AR/VR technology.

Figure 2. Research team members record exact 3D coordinates using a total station surveying instrument (upper left), capture video of an excavation using a digital camera (lower left), and scan a portion of the dig site using the Cyrax range scanner (right).

2.2. Collaborative AR/VR Systems

Since there has been much work on collaborative VR applications we focus our review on AR/VR systems that use tracked, rather than stationary, displays and those that respond to multimodal interaction. The most common, yet expensive, solution to a collaborative VR experience is to use a CAVE(-like) display system. Those systems, as well as responsive workbenches [16], allow for stereoscopic viewing, but present all users with only a single user's correct viewpoint. This makes it impossible to provide a perspectively correct view of the environment to more than one user at the time.

Several groups have tried to address this limitation. Agrawala et al. [2] have demonstrated a two-user responsive workbench in which four separate frame buffers (one for each of the two users' eyes) are time sequenced, thereby allowing up to two users to have a correct stereoscopic view of the environment at half the regular frame rate. The IllusionHole of Kitamura et al. [15] overlays a mask with hole over a responsive workbench, giving each of a small set of users an independent stereoscopic view in the portion of the display that user sees through the hole. However, major limitations still exist with both technologies, including the limited number of users.

Efforts have been made recently towards tele-immersive collaboration environments. Raskar et al. [22] present an approach to virtually join two office spaces using large wall-projected displays and computer vision techniques. Broll at al. [6] have created "The Virtual Round Table," where multiple people can collaborate using real video feeds in a virtual environment.

AR merges the real world with superimposed virtual images, combining the advantages of both real and virtual environments. Billinghurst and Kato [4] pointed out the benefits of AR in collaborative settings and how such systems decrease the cognitive and functional load on the user. Several collaborative AR systems have used either see-through head mounted displays [4,24], tracked hand-held LCD displays [23], or both [7].

Hua et al. [13] have recently presented another effort to create a rich AR environment. They use a relatively wide field-of-view projective see-through head-worn display, and wall surfaces and an interactive workbench that are covered with retroreflective material, to facilitate 3D navigation tasks in an immersive environment. In their system, multimodal-based interaction is limited to 2.5D interaction with projected surfaces and physical markers on the responsive workbench.

3. Application Scenario

In our prototype, two users wearing see-through, head-worn displays and tracked gloves can explore a portion of the virtual dig site. They are able to view a 3D representation of the terrain as either a textured point cloud or a textured mesh. Small archaeological finds are placed around the dig site at the exact locations of their discovery, each labeled with its name and description. Users can navigate the site through the use of a multi-touch, multi-user, interactive table surface (MERL DiamondTouch table) on which they can see a bird's eye view of the entire site, select objects for further inspection, view additional multimedia about the dig site, such as movies and photographs, and adjust personalized view settings in their head-worn displays.

Objects can be viewed and inspected both in 2D on the DiamondTouch table and in 3D through the head-worn display. While in the VR environment, an object can be selected by grabbing it and "throwing" it on the DiamondTouch table, or by specifying it multimodally at a distance through speech and gesture. Once an object is selected, it virtually appears next to the DiamondTouch table for 3D visualization, while additional 2D information and multimedia about that object are presented on the table. As shown in Figure 1, a high-resolution hand-held display can be used as a "magic lens" [5] to view additional data and detail on a portion of the DiamondTouch table.

4. Collaborative AR Environment

The system presented here builds upon our previous multimodal interaction work [14,20]. We have tried to facilitate as many interaction modalities and fuse them together, giving the user a choice of which one to use at certain situations. Our paper presents a working prototype and should be considered a work in progress.

4.1. Modular Approach

To facilitate collaborative visualization of archaeological data, we have designed a modular head-tracked AR/VR environment, augmented with a front-projected multi-user multi-touch table and a rear-projected wall display.

As shown in Figure 3, our system consists of a set of AR/VR visualization modules (AR/VR modules), a DiamondTouch multi-user table module (DT module), and a rear-projected wall display (WALL module). The AR module consists of a tracked see-through head-worn display (Sony LDI-D100B), a tracked glove (Essential Reality P5 glove [9]), and a stereo headphones and microphone, supported by speech-recognition software (IBM ViaVoice 10), and an overhead six–degree-of-freedom tracking infrastructure (InterSense IS-900). The DT module consists of the MERL DiamondTouch [18] table, a projector (InFocus Proxima x350) and a hand-held display (Fujitsu Stylistic LT C-500). The WALL module displays information on our rear-projected wall display. Each module is currently running on a separate dedicated PC (dual AMD Athlon MP 2.0, 1GB RAM).

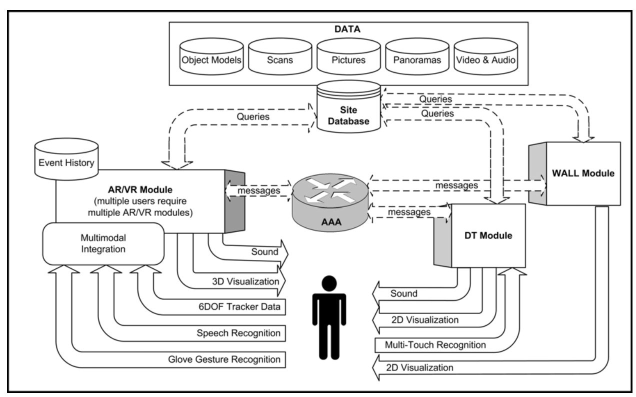

Figure 3. System architecture. A separate AR/VR module is needed for each user.

All collaboration and communication is conducted via simple message passing through a publish-and-subscribe message board system. We are using the Adaptive Agent Architecture (AAA) [17] to facilitate easy connection, discovery and communication management of our modules. All modules can function completely independently of each other.

To reduce the number of messages and their size, all users have direct access to the same database containing all available archaeological materials from the dig site, and only the most relevant information is communicated. For example, if an AR user selects an object, only the distinct object ID will be broadcasted, allowing all other modules to retrieve all necessary information about that object directly from the database.

The number of simultaneous users depends only on the number of available AR modules and is physically capped by the throughput of the AAA and available hardware. Furthermore, users do not need to be co-located, since the modularity of our approach allows for remote collaboration; however, in this paper we describe only co-located collaboration that also uses our DT module.

4.2. Multimodal Interaction

We are developing hybrid interaction techniques for exploring our virtual and augmented environment. Users have a variety of means for interacting with the 3D environment, including combinations of gestures and speech, and four buttons attached to the glove. Since no modification of the environment is allowed, users' interaction is primarily focused on selection and inspection of objects, terrain, and multimedia data. Object selection is the most frequently used 3D interaction technique, and we provide several ways to accomplish it. The user can walk towards an object and grab it or point at it in the distance using our SenseShapes selection tools [20], as shown in Figure 4. Selection at a distance is prone to ambiguity, especially when many similar objects are near each other, and SenseShapes assist in their disambiguation.

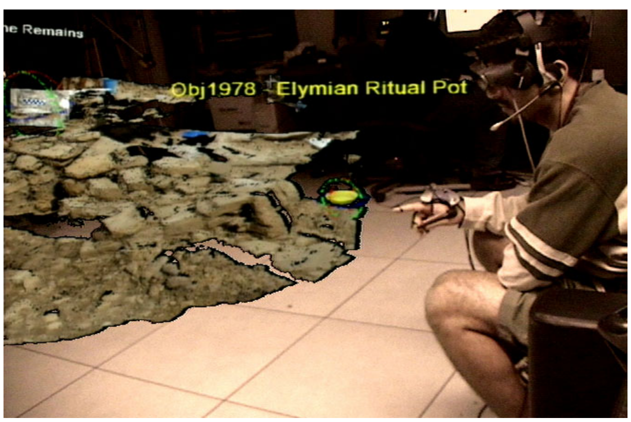

Figure 4. User's view of a section of the site in AR. The other user, 3D terrain model, and objects in this section are visible. Objects are represented either with a 3D model or with a picture if the model is not available.

Figure 5. User inspecting objects in his "virtual pack."

SenseShapes are 3D selection volumes that represent the regions of interest and are attached to a tracked portion of the user's body (e.g., the head and hand, in our case). SenseShapes keep an event history of all the objects that intersect them and provide a way to query and obtain statistics on those objects. For example, it is possible to request "the objects that were in the picking cone during interval *T*" and SenseShapes will return the list of those objects, as well as useful statistics about them to facilitate correct multimodal interpretation. The statistics include information about time (how long was the object in the volume), stability (how many times did the object enter and leave the volume), distance (how far was the object from the user), and visibility (how much of the object was visible during that time frame). These statistics, combined with our gesture recognizer and speech recognizer, form our multimodal interaction framework. Our glove-based gesture recognizer currently supports three distinct gestures (point, grab, and thumbs-up).

Once an object of interest is selected, the user can save her selection in a "virtual pack" (shown in Figure 5) by pressing one of the buttons on her glove. A virtual copy of the object is then placed in the pack that surrounds the user. Using different buttons on the glove, the user can easily store, inspect, and remove objects from the pack. The "virtual pack" metaphor is similar in notion to work Pierce et al. [21], who demonstrated widgets for storing and retrieving objects in 3D environments. We envision that this pack metaphor will be the primary collection method in our system. The pack is personal and private to each user; however the user can make the pack visible to other users via speech command or transfer the contents of the pack to the DT module for further object inspection.

Since the users in our system are being continuously tracked and recorded, normal human-to-human communication can be a source of many unwanted errors. To facilitate a high level of human-to-human interaction among users it is necessary to be able to control all the modalities to prevent unwanted actions from executing. In our system, each modality can be turned on or off. We tried to place each "modality switch" into an intuitive logical location: The glove can be turned on/off with a simple button press on the glove itself, while speech interpretation can be engaged and disengaged by saying a voice command.

4.3. Multi-touch, Multi-user Projection Table

Several users can simultaneously interact with a MERL DiamondTouch projection table to collaboratively navigate through the archaeological site. (The table distinguishes simultaneous interactions by multiple users, who are each in contact with a different conductive pad, typically by sitting on it.) For example, if a user selects an object that appears on the 2D table, as shown in Figure 6, all users can then view information about the selection, as well as add additional annotations. Hand-held computer displays resting on and tracked by the table can serve as physical "magic lenses"; these can provide better resolution (pixel density) than the imagery projected on the table, while also supporting customization based on the particular user manipulating the hand-held display (recognized by their touch). In essence, they augment the table so that it acts as a multi-user focus-plus-context display with multiple, tangibly-controlled foci. Thus, complementary information can be displayed on the projected table display, on tracked hand-held displays, or on tracked see-through head-worn displays.

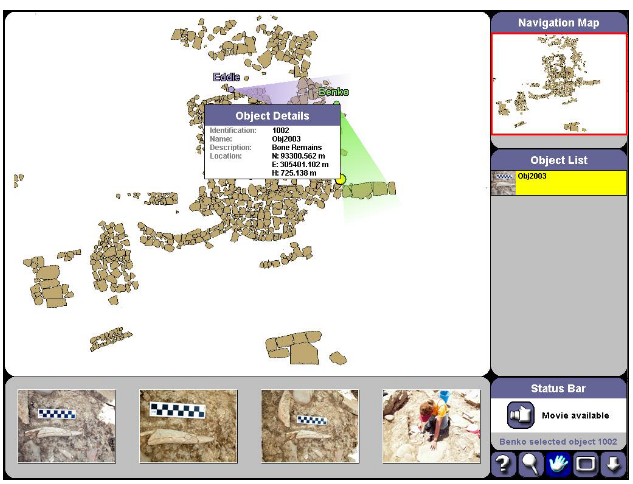

Figure 6. The DT module user interface, showing users being tracked in the 3D environment and information about the currently selected object.

4.4. Putting it all Together

We have described several ways in which multiple users can visualize archaeological data. However, all displays are not suitable for each interaction method, and all methods are not equally suitable for different scenarios. For example, speech-driven interaction is much more appropriate when users are not communicating directly with each other or when only one user is interacting with the system, whereas in collaboration among multiple users usually speech may be better reserved for human-to-human communication, with gestural interaction used for communicating with the system.

We use four kinds of displays in our system: tracked, see-through, head-worn, 3D displays; a front-projected multi-touch, multi-user 2D table; hand-held 2D displays for viewing higher-resolution personalized data on portions of the projected table surface; and a wall-sized rear-projected display.

These displays vary in the amount of personalized data they can present. All users share the displayed information on the interactive, multi-touch table, as well as the hand-held displays. In addition, the control of both displays is also shared. In contrast, the 3D head-worn displays are capable of displaying personalized views of the environment. In addition to head-tracking, control of what the user sees on their head-worn display is accomplished through the multi-touch table or by multimodal interaction. Since the DiamondTouch table can distinguish between different simultaneous users, it can be used to control the material presented on multiple head-worn displays.

5. Discussion

Our system is a work-in-progress and we are still in the early stages of experimenting with our prototype application. We discuss here some of the potential advantages of our approach and its current limitations.

5.1. Advantages

Multiple interaction techniques. We have incorporated multiple interaction techniques in our system, including speech, gesture, 3D multimodal interaction, and 2D interaction on a multi-user table. Not all modalities are adequate or even desired at all times for all tasks, and therefore careful management of the use of these modalities can produce the best results. Additionally, since human-to-human interaction is important, our system is built to facilitate it.

3D multimodal interaction. We use the SenseShapes techniques from our previous work to facilitate natural interaction in the 3D environment and minimize possible ambiguities in the selection of objects.

Multiple users. Multiple simultaneous users are supported, limited only by available hardware and the throughput of the message board system.

Hybrid visualization space. Our system provides both 2D and 3D immersive displays, and therefore provides a range of choices for visualizing data.

Personalization. Each user's view can be customized, based on their specific needs. Users can also share some of their personal information with other users; for example, one user can request to see objects collected in another user's virtual pack.

Remote and co-located collaboration. Although we present only co-located examples in this paper, our system can also be configured for remote collaboration by representing remote users as avatars.

5.2. Limitations

Data Classification and contextualization. We need a completely classified database of archaeological data, together with linguistic keywords for better speech recognition and multimodal interpretation.

Limited test data. We currently have a small number of actual scanned objects, as well as a limited set of scans of the dig site, acquired during our relatively short ten days on-site. To provide for a more complete immersive experience, much more data must be acquired. Because of limitations of the Cyrax scanner that our team used, we would need a separate scanner designed for scanning small objects to capture artifacts and finds.

DiamondTouch. Although the DiamondTouch table can detect simultaneous independent touches by multiple users, it has some limitations. Since a grid of charge-sensitive antennas detects the users' touches, specific 2D points are cannot be known, but rather, the charged vertical and horizontal antennas produce intersection points. Therefore, it is typically used to detect selection boxes defined by the axis-aligned bounds of a given user's touches.

Head-worn displays. Our head-worn displays' field of view (30° horizontal) is far too small, and their obstruction of the user's eyes makes face-to-face communication difficult.

Tracking. Our IS900 tracking system supports only four trackers, and each user requires two trackers (head and hand); thus, we are currently limited to only two co-located users.

6. Conclusions and Future Work

We have presented a VR/AR system allowing archaeologists to collaboratively discuss and analyze an excavated dig site. We allow multiple users to walk around the virtual site, embedded with both 3D and 2D multimedia data, and to multimodally inquire about interesting finds in-situ. We also provide a collaborative table surface allowing multiple users to simultaneously touch and further inquire about the site.

In the future we hope to further extend this prototype with more gestural control of the environment and the pack. We hope to start work on developing a dynamic decision-making framework that would decide which display (combination) to use when given specific data characteristics, user preferences and user capabilities. Furthermore, we plan to explore ways to personalize the user interface and the user experience in our system.

Finally, we hope to allow for remote collaboration and do a user study with archaeologists. Since many excavations are a joint effort by multiple archaeologists from varying geographic locations, potentially much post-excavation analysis and interpretation is done remotely, which makes communication quite difficult. Our system has the potential to allow multiple archaeologists to visualize and navigate the same virtual site remotely, communicating through voice and gestures.

Acknowledgements

We wish to thank Ian Morris and the Stanford Archaeology Center for allowing us to participate in Monte Polizzo dig, Peter Allen, Alejandro Troccoli, and Ben Smith for the scanning data, Lynn Meskell and James Conlon for their archaeological expertise, Sajid Sadi and Avinanindra Utukuri for support with the P5 glove, and Shezan Baig for his work on surface meshing. This work is funded in part by NSF Grants IIS-01-21239 and IIS-00-82961, Office of Naval Research Contracts N00014-99-1-0394, N00014-99-1-0683, and N00014-99-1-0249, and gifts from Mitsubishi Electric Research Labs and Microsoft.

References

[1] Acevedo, D., Vote, E., Laidlaw, D.H., and Joukowsky, M.S. "Archaeological Data Visualization in VR: Analysis of Lamp Finds at the Great Temple of Petra, a Case Study." In *Proc. IEEE Visualization 2001*. San Diego, California. October 2001, pages 493–496.

[2] Agrawala, M., Beers, A., Fröhlich, B., Hanrahan, P., McDowall, I., and Bolas, M. "The two-user responsive workbench: Support for collaboration through individual views of a shared space." In *Computer Graphics (SIGGRAPH '97 Proceedings)*, 1997, pages 327–332.

[3] Autodesk AutoCAD. http://www.autodesk.com.

[4] Billinghurst, M., and Kato, H., "Collaborative mixed reality" in *Mixed reality: Merging real and virtual worlds*, Y. Ohata and H. Tamura (Eds.), Co-published by Ohmsha and Springer-Verlag, 1999, 261–284.

[5] Bier, E.A., Stone, M.A., Pier, K., Buxton, W., and DeRose, T.D. "Toolglass and magic lenses: The see-through interface." In *Proc. SIGGRAPH '93*, 1993. 73–80.

[6] Broll, W., Meier, E., and Schardt, T. "The Virtual Round table—A collaborative augmented multi-user environment." In *Proc. Collaborative Virtual Environments* (*CVE 2000*)*,* San Francisco, CA. 2000.

[7] Butz, A., Höllerer, T., Feiner, S., MacIntyre, B., and Beshers, C. "Enveloping users and computers in a collaborative 3D augmented reality." In *Proc. IWAR '99*

(IEEE and ACM Int. Workshop on Augmented Reality), San Francisco, CA, October 20–21, 1999, 35–44.

[8] Cruz-Neira, C., Sandin, D. J., and DeFanti, T.A. "Surround-screen projection-based virtual reality: The design and implementation of the CAVE." In *Proc. SIGGRAPH '93*. 1993. 135–142.

[9] Essential Reality Inc. http://www.essentialreality.com.

[10] ESRI ArcGIS. http:// www.esri.com.

[11] INTRASIS, The Swedish National Heritage Board, Archaeological Excavations Department and LandFocus IS AB. http://www.raa.se/uv/intrasis/eng/index.htm.

[12] Gaitarzes, A. Christopoulos, D., and Roussou, M. "Reviving the past: Cultural Heritage meets virtual reality." In *Proc. Conf. on Virtual Reality, Archeology, and Cultural Heritage.* 2001. 103–110.

[13] Hua, H., Brown, L.D., Gao, C., Ahuja, N. "A new collaborative infrastructure: SCAPE." In *Proc. IEEE Virtual Reality (VR '03),* 2003.

[14] Kaiser, E., Olwal, A., McGee, D., Benko, H., Corradini, A., Li, X., Cohen, P., and Feiner, S. "Mutual disambiguation of 3D multimodal interaction in augmented and virtual reality." To appear in *Proc. ICMI 2003 (Fifth Int. Conf. on Multimodal Interfaces*), Vancouver, BC, November 5–7, 2003.

[15] Kitamura, Y., Konishi, T., Yamamoto, S., and Kishino, F. "Interactive stereoscopic display for three or more users." In *Proc. SIGGRAPH '01*. 2001. 231–240.

[16] Kröger, W. and Fröhlich, B., "The responsive workbench." *IEEE Computer Graphics and Applications*. 1994. 12–15.

[17] Kumar, S., Cohen, P. R. and Levesque, H. J. "The Adaptive Agent Architecture: Achieving Fault-Tolerance Using Persistent Broker Teams." In *Proc. Int. Conf. on Multi-Agent Systems*, July 2000, pages 159–166.

[18] MERL DiamondTouch Table. http://www.merl.com/projects/DiamondTouch/

[19] Monte Polizzo, The Stanford Project, Stanford Archaeology Center, Stanford, CA. http://archaeology.stanford.edu/MountPolizzo/project.htm

[20] Olwal, A., Benko, H., and Feiner, S. "SenseShapes: Using statistical geometry for object selection in a multimodal augmented reality system." In *Proc. ISMAR 2003* (*IEEE and ACM Int. Symp. on Mixed and Augmented Reality*), Tokyo, Japan, October 7–10, 2003, pages 300–301.

[21] Pierce, J.S., Conway, M., Dantzich, M., Robertson, G. "Toolspaces and glances: storing, accessing, and retrieving objects in 3D desktop applications." In *Proc. Symp. on Interactive 3D graphics* (*I3D '99*). 1999. 163–168.

[22] Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., and Fuchs, H. "The office of the future: A unified approach to image-based modeling and spatially immersive displays." In *Proc. SIGGRAPH '98.* 1998. 179–188.

[23] Rekimoto, J. "Transvision: a hand-held augmented reality system for collaborative design." In *Proc. Virtual Systems and Multimedia (VSMM '96)*. Gifu, Japan, 1996.

[24] Szalavari, Z., Schmalstieg, D., Fuhrmann, A., and Gervautz, M. "Studierstube: an environment for collaboration in augmented reality." *Virtual Reality, 3(1),* 37–48. 1998.