

Collaborative Mixed Reality Visualization of an Archaeological Excavation

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

Abstract

We present VITA (*Visual Interaction Tool for Archaeology*), an experimental collaborative mixed reality system for offsite visualization of an archaeological dig. Our system allows multiple users to visualize the dig site in a mixed reality environment in which tracked, see-through, head-worn displays are combined with a multi-user, multi-touch, projected table surface, a large screen display, and tracked hand-held displays. We focus on augmenting existing archaeological analysis methods with new ways to organize, visualize, and combine the standard 2D information available from an excavation (drawings, pictures, and notes) with textured, laser range-scanned 3D models of objects and the site itself. Users can combine speech, touch, and 3D hand gestures to interact multimodally with the environment. Preliminary user tests were conducted with archaeology researchers and students, and their feedback is presented here.

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 ongoing archaeological excavation. The dig site was on top of Monte Polizzo in western Sicily, where a team of archaeologists [30] have been excavating an Elymian acropolis built between the 6th and 4th centuries B.C. Over a span of ten days, we used a 3D laser range scanner, a total station surveying system, and digital video and still cameras to collect 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 [3].



Figure 1. Two users simultaneously collaborate in VITA. While one user is inspecting the 3D virtual model of the ceramic vessel above the table comparing it with the high-resolution image on the screen, the second user is looking at the 3D miniature terrain model next to the table. All AR images in this paper are captured through a live tracked video see-through display.

The excavation of an archaeological site is by nature a destructive, and often physically unreconstructable, process. Soil layers are removed to uncover the layers beneath, small finds are taken away from the site for conservation and future analysis, and previously excavated structures are sometimes completely deconstructed to uncover even earlier structures. Detailed recording of each stage of the excavation is crucial when attempting to reconstruct the previously excavated layers. However, even though archaeologists already try to carefully record every part of an excavation, each team member usually works on one particular section of a site, thus making it difficult to analyze, contextualize, and connect the resulting data across the entire site.

Based on our interviews with archaeologists, we determined that much of the typical post-excavation analysis, interpretation, report writing, and additional research they perform 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. The archaeologists envision using this space both for 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 should be able to 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.

To address these needs, we present a collaborative mixed reality system, VITA (*Visual Interaction Tool for Archaeology*), shown in Figure 1, which was designed in collaboration with a group of archaeologists, to help users, ranging from interested novices to experienced archaeologists, visualize and explore excavation results off-site. We have focused on complementing standard archaeological recorded data, which is mostly 2D in nature (drawings, pictures, and notes), with a variety of acquired data (3D panoramic images, 3D models of small finds and of the site itself at various stages of excavation, video, and ambient audio), and making it all available in one seamless collaborative environment.

2. Related Work

Because we are developing for the specific domain of archaeology, 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 and guidelines for recording the state of the dig site during an excavation, their main focus is to record and archive the data, rather than to visualize it (e.g., [4, 17]).

To visualize their data, most archaeologists currently rely on geographic information systems (GIS), such as the ESRI ArcGIS [5] suite of software. Recently, INTRASIS [19] has extended some capabilities of standard GIS systems and functions as a plug-in for ArcView GIS. Furthermore, standard computer-aided design (CAD) systems, such as AutoCAD [6], are often used for modeling and reconstruction, but 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 [13]. 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. [16] present various VR setups, ranging from an Imersadesk™-based interactive plane to a CAVE-based environment, for visualizing temples and public buildings in ancient Greece.

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 employ multimodal interaction. The most common, yet expensive, way to provide a collaborative VR experience is to use a CAVE(-like) [13] display system. Those systems, as well as responsive workbenches [22], support 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 both head-tracked users to have a correct stereoscopic view of the environment. but at half the regular frame rate. The IllusionHole of Kitamura et al. [21] overlays a mask with hole over a responsive workbench, giving each of a small set of head-tracked users an independent stereoscopic view in the relatively small portion of the display visible to that user alone through the hole.

Several efforts have targeted the development of tele-immersive collaboration environments. Raskar et al. [28] present an approach to join two office spaces virtually by using large wall-projected displays and computer vision techniques. Broll et al. [11] 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 [10] 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-worn displays [10, 32], tracked hand-held LCD displays [29], or both [12, 24].

Hua et al. [18] have recently presented another effort to create a rich AR environment. They use relatively wide field-of-view projective see-through head-worn displays, together with wall surfaces and an interactive workbench that are covered with retroreflective material, to facilitate 3D navigation tasks in an immersive environment. However, in their system, multimodal-based interaction is limited to 2.5D interaction with projected surfaces and physical markers on the workbench, and a user can only see virtual material where they have a clear view of a retroreflective surface.

3. VITA System

We designed the VITA system to make it possible for archaeologists to collaborate in a shared hybrid environment space, wearing tracked, head-worn displays to visualize 3D terrain data and embedded multimedia. In addition to their head-worn displays, users also have available a large, high-resolution display, a tracked handheld display, and a multi-user, multi-touch, projected table surface. Each user can use a tracked glove, speech commands, and the multi-touch sensitive surface to interact multimodally with the system and collaborate by jointly navigating, searching, and viewing data.

Exploration of the environment is supported by two basic modes: world-in-miniature mode and life-size-world mode. The *world-in-miniature* mode [31] (shown in Figure 1) presents the user with a small-scale virtual model of the acropolis at Monte Polizzo, fixed on the physical tabletop next to the touch-sensitive projected surface onto which we project our 2D user interface. The 2D display shows a modified *Harris Matrix* [17], a widely used archaeological diagram representation of temporal relationships in the excavation of all the archaeological layers on-site (shown in Figure 2). For example, an archaeological layer could be 10 cm of soil representing the ground level of a particular time period (e.g., mid 4th century B.C.). By selecting portions of the Harris Matrix, users can choose to visualize various data relevant to the archaeological layers and finds that they are currently exploring. At the same time, the currently selected layers and finds appear in the scaled 3D model, making apparent the spatial relationships between them. In addition, a large, high-resolution screen is available for viewing documents (e.g., handwritten field notes or photographs) and a high-resolution hand-held display can be used as a “magic lens” [9] to view detailed 2D data on a portion of the touch-sensitive table.

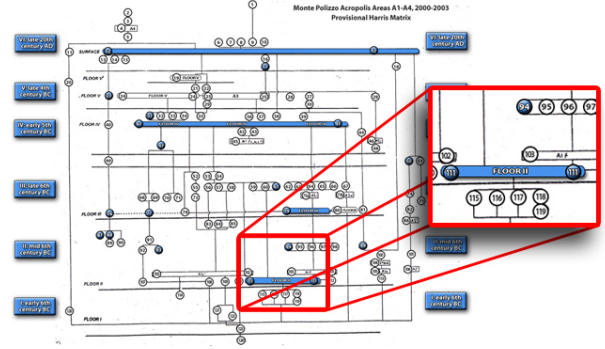


Figure 2. The enhanced Harris Matrix conveys the temporal relationship of the excavated layers. All blue areas are hotspot links to more data about that particular layer. (The outlined detail view was added for clarity.)



Figure 3. Real vs. virtual: (a) Real image of a portion of the excavated structure taken at the site. (b) User exploring the same section of the site in life-size-world mode of VITA; the other user, 3D terrain model, and several bone finds in this section are visible.

In *life-size-world* mode, the textured, meshed model of the Monte Polizzo acropolis is displayed in the user's head-worn display, covering an area of roughly 10×10 meters, as shown in Figure 3(b). Users wearing see-through head-worn displays and tracked gloves can walk around the site at its actual size. To recreate the experience of being on top of Monte Polizzo, users can request the system to display a surrounding panorama to provide a completely immersive experience. Using a tracked glove, a user can examine archaeological finds that are situated in the model at the exact locations of their discovery, each labeled with its name and description. The finds are represented in the model as colored diamonds (similar to the pyramids of [1]) where color is based on the object type; for example, ceramic objects are red and wooden objects are blue. (In an earlier version, we positioned photorealistic 3D models of these objects in the environment, but in most cases, due to their color similarity with the surrounding soil, and the drastically varying sizes of the objects, life-sized photorealistic representations were very difficult to visualize in situ.)

3.1. System Architecture

To facilitate collaborative visualization of archaeological data, as well as allow the use of a variety of different displays, we have designed a modular environment. As shown in Figure 4, our system consists of a set of AR visualization modules (AR modules), a multi-user multi-touch table module (DT module), a large high-resolution display (SCREEN module), and a handheld high-resolution display for use with the table surface (HANDHELD module).

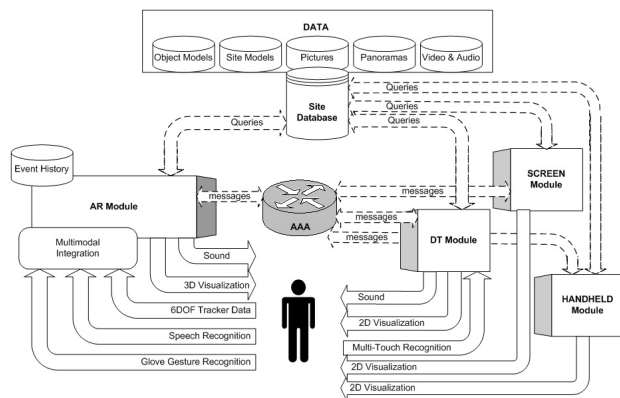


Figure 4. VITA system architecture. A separate AR module is needed for each user.

The AR module consists of a tracked, see-through, head-worn display (Sony LDI-D100B); a tracked glove (Essential Reality P5 glove [26]); stereo headphones and microphone, supported by speech-recognition software (IBM

ViaVoice 10); and an overhead six-degree-of-freedom tracking infrastructure (InterSense IS-900). A dedicated computer (Dual Athlon MP 2.0, 1GB RAM) runs one AR module per user. Due to current hardware constraints, an AR module requires a large number of connecting wires to power and control all the devices (as seen in Figure 5), making it fairly cumbersome for the user to move around freely.

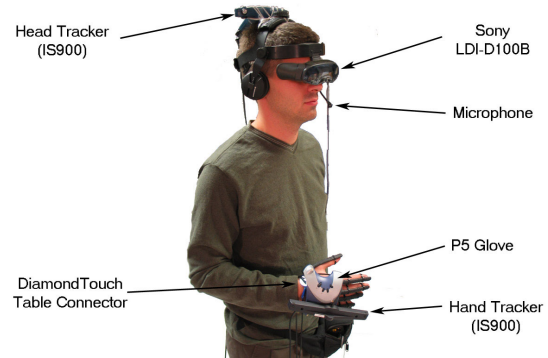


Figure 5. User's AR interaction devices.

The DT module consists of the front projected (InFocus Proxima x350, 1024×768 resolution) MERL Diamond-Touch [14] table capable of detecting multiple touches and disambiguating among multiple users. The SCREEN module displays information on a 24" LCD monitor (Samsung 240T, 1920×1200 resolution). The HANDHELD Module, powered by a TabletPC (HP TC1100, 1024×768 resolution), is used to provide higher resolution imagery when needed on the DiamondTouch surface.

3.2. Module Communication

All communication between modules is conducted via simple message passing through a publish-and-subscribe message board system. We are using the Adaptive Agent Architecture (AAA) [23] to facilitate easy connection, discovery, and communication management of our modules. While the best functionality is achieved with all the modules present, each one can act as a completely independent, self-contained unit.

To reduce the number of messages and their size, all modules 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 broadcast, allowing all other modules to retrieve all necessary information about that object directly from the database. Thus, 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. Further-

more, 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.

4. Interaction in VITA

The VITA system incorporates a variety of modalities, described below, where each one is used to best serve a particular task, but all work together to form an integrated system. Tabletop interaction (Section 4.1) allows 2D navigation of all objects and their relations to the various layers of the site. Used in conjunction with the table, hybrid gestures (Section 4.2) [8] manipulate 3D models of objects, while the handheld widget (Section 4.3) creates a movable high-resolution portal within the tabletop. Finally, while navigating the life-sized site model, users can make use of multimodal interaction (Section 4.4) [20, 25] in conjunction with our VirtualTray (Section 4.5).

4.1. Tabletop Interaction

Several users can simultaneously interact with the MERL DiamondTouch projection table to collaboratively navigate through the dig site. The table distinguishes simultaneous interactions by multiple users because each is in contact with a different conductive pad (e.g., by sitting on or touching it). Figure 5 shows several layer windows, including a selected one painted cyan, with its highlighted location in the Harris matrix. Although temporal relationships of layers and objects are viewed on the table's surface, all scanned, high-resolution documents pertaining to a particular layer are viewed on the high-resolution monitor to avoid scaling and resizing on the lower-resolution projected surface.

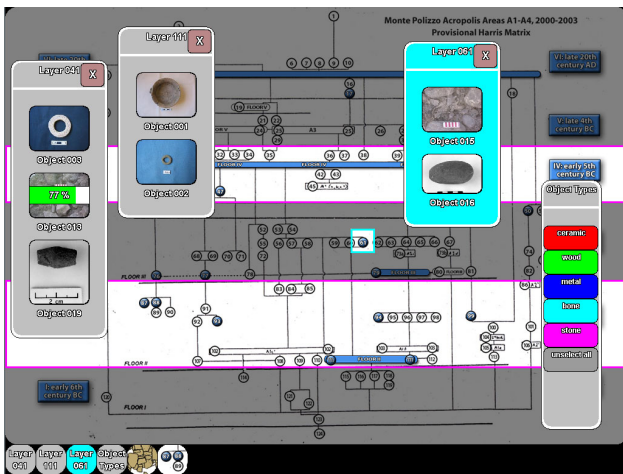


Figure 5. The DT module's user interface, showing information about several layers and objects selected from the Harris Matrix. The currently selected layer is painted cyan and is highlighted in the background Harris Matrix, thus visually showing the context.

4.2. Hybrid Gestures

Integrated into the VITA system is a set of interaction techniques designed specifically for hybrid user interfaces that use 2D touch-sensitive displays and 3D AR user interfaces [8]. The motivation for these techniques comes from the archaeologists' need to visualize 3D representations of archaeological finds in context with the 2D data for those same objects. The interactions are based on synchronized 2D and 3D gestures that facilitate a seamless transition of data across dimensions. Single-handed gestures such as cross-dimensional pull (shown in Figure 6) allow users to transition objects between the 2D and 3D environments. We also facilitate two-handed interaction combinations to adjust multiple independent transformation parameters of an object simultaneously, such as a combination of scale and 3D rotation.



Figure 6. Using the cross-dimensional pull gesture to replace a 2D image of a pottery find on the table with a 3D model above the table.

4.3. Handheld Widget

Handheld computer displays resting on and tracked by the table can serve as physical "magic lenses" [9]. They can provide better resolution (higher pixel density) than the imagery projected on the table (see Figure 7), while also supporting customization based on the particular user manipulating the hand-held display (recognized by their touch transmitted by a conductive handheld display case).



Figure 7. Handheld widget is used on the Diamond-Touch table to show high-resolution version of the portion of the 2D map that it covers.

The projected image is suppressed in the area in which a hand-held display is located. In essence, the handheld displays augment the table so that it acts as a multi-user focus-plus-context display with multiple, movable, tangibly-controlled foci. This builds on two pieces of earlier work that embedded a single immobile focus within a lower-resolution context. Both earlier approaches used a stationary flat-panel focus, viewed within a context presented in one case on a head-tracked, see-through, head-worn display [15] and in the other case on a stationary projected surround [7]. Thus, complementary information can be displayed on the projected table display, on tracked hand-held displays, or on tracked see-through head-worn displays.

4.4. 3D Multimodal Interaction

In life-size-mode exploration, the users are often not sufficiently close to the table to use its 2D user interface to navigate the world. Rather than returning to the table for each adjustment, we have implemented a series of speech- and gesture-driven interactions to make it possible to navigate in place.

In line with trying to present the dig site as close as possible to the way it was during excavation, modifying the environment is not allowed. Therefore, a user's interactions are primarily focused on selecting and inspecting 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 toward an object and grab it, or point at it in the distance using our SenseShapes selection tools [25]. Selection at a distance is prone to ambiguity, especially when many similar objects are near each other, and SenseShapes assist in their disambiguation.

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 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 the object was in the volume), stability (how many times the object entered and left the volume), distance (how far the object was from the user), and visibility (how much of the object was visible during the interval). 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). In addition to multimodal selec-

tion, simple speech commands can be employed to customize and adjust visualization parameters.

4.5. VirtualTray Widget

When exploring in life-size-mode, we noticed that the users often wanted to collect several objects from the model based on where they were and then inspect them in more detail once they returned to the DT Module. To facilitate this, we implemented a 3D widget we call the *VirtualTray*, shown in Figure 8.



Figure 8. User inspecting objects in his VirtualTray.

Once an object of interest is selected, the user can save her selection in their VirtualTray by pressing one of the buttons on her glove. A virtual copy of the object is then placed in an annular "tray" that surrounds the user. Using different buttons on the glove, the user can easily store, inspect, and remove objects from the tray. The VirtualTray metaphor is somewhat reminiscent of a "lazy susan" and is similar in notion to work by Pierce et al. [27], who demonstrated widgets for storing and retrieving objects in 3D environments. Each user has their own VirtualTray, which is invisible to other users; however a user can choose to transfer the contents of her VirtualTray to the DT Module for further object inspection. At that point all of a user's collected objects are publicly displayed on the tabletop surface.

5. Design Considerations

We designed VITA with experienced archeologists as our target users. Many of our ideas and design decisions grew out of discussions with archeologists on our team and the focus of our development was centered on four key ideas: (a) make a collaborative visualization system, (b) provide a simpler way to explore all "standard" archaeological data, (c) incorporate a widely accepted analysis paradigm, such as Harris matrix, into the environment, and (d) provide novel visualization and interaction ways to explore the excavation results. In addition, our target user was not necessarily computer knowledgeable; therefore, we tried to design our interactions to be intuitive and relatively simple.

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 in a multimodal system. 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” in an intuitive logical location; for example, a glove can be turned on/off with a simple button press on the glove itself, while speech interpretation can be engaged and disengaged with a voice command.

The necessity of requiring multiple displays became apparent very early in our development when we realized that we were dealing with a wide variety of data, each with different size, resolution and spatiality (2D or 3D) requirements. A guiding principle was to use the most appropriate display available to present each kind of data given the data medium. For example, a textured 3D model may be best viewed in a tracked, head-worn, stereoscopic display, while high-resolution scanned field notes are best displayed on a large high-resolution monitor; the video might be displayed on the large screen, while spatially varying layer information is best viewed on the tracked handheld. In addition, this setup allowed us to minimize data clutter and avoid much moving and resizing by forcing some data to appear on a dedicated surface.

We use four kinds of displays in our system: tracked, see-through, head-worn, 3D displays; a single front-projected tabletop display; hand-held displays; and a single large high-resolution flat-panel display. While all users share the displayed information and the control of the tabletop, handheld and large size display, 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 table can distinguish among different simultaneous users, it can be used to control the material presented on multiple head-worn displays. Currently, in our system, users can privately visualize the VirtualTray, as well as some of the 3D object models, by using private hybrid gestures. However, we are considering the possibilities of providing two different users (for example, a ceramics researcher and a metallurgy researcher) with personalized views of the same environment in their head-worn displays.

6. User Experience

To perform an early evaluation of the effectiveness of VITA, we invited three archaeologists and three archae-

ology graduate students to try out the system informally. The archaeologists were extremely familiar with the Monte Polizzo dig site, having been active participants in previous years' excavations, while the students, although familiar with the general practice of archaeological excavation, were completely new to this particular dig site. Each user was given an hour-long introduction to the system, in which they learned about its capabilities and constraints. They practiced selecting objects in 3D, performing hybrid gestures, and manipulating data on the DiamondTouch table. During that hour, each user was accompanied by at least one of the system developers, who guided them through different aspects of the system. The current two-user tracking limitation in VITA forced us to constrain collaboration to two people at a time.

Initial user reactions were very positive, primarily including comments regarding the 3D visualization of the excavated circular structure, stating it was both beneficial and complementary to the 2D context sheets, drawings, and diagrams that are traditionally available post-excavation. Furthermore, manipulation (e.g., scale and rotation) of the 3D model was very useful, allowing the archaeologists to capture an arbitrary viewpoint with little effort. Visualizing the life-sized model was important to see the details of each stone and the intersections of walls, while visualizing the world-in-miniature was also important to obtain a bird's eye view of the model in context to the rest of the dig site. All users pointed out that they thought the major benefit of our system was the ability to connect the temporal relationships of excavated objects (in the Harris matrix) with their 3D spatial relationships, all while providing contextual 2D information for those objects. VITA facilitates this capability in seconds, when normally this task can take hours to complete.

Criticism of the system focused mainly on missing features that could not be implemented because of the unavailability of necessary data. In particular, users wished to see how layers were spatially connected over time through a time-lapse visualization. This problem could have been resolved during the data acquisition phase several months earlier on Monte Polizzo, had we only been able to acquire more frequent scans over a much longer period of time. Collaborating users were also frustrated when one user would not know where another was looking in the augmented environment. A major contributing factor was the inability to see the other user's eyes because of occlusion by the dark filters built into the head-worn displays; however the use of a virtual laser pointer could possibly have aided in lessening this frustration. Also, as expected, users complained about the discomfort of wearing the interaction devices for an extended period of time, saying it was too heavy and had too many wires,

which often became tangled when walking around the virtual dig site.

6.1. Evaluation of a Learning Tool Scenario

In addition to the general usage of VITA, we informally investigated how effective the system might be as a collaborative learning tool. We asked one of the archaeologists to use VITA as a tool to teach students about specific aspects of the Monte Polizzo site. In this scenario, the archaeologist was given a half hour to teach a student about the site, finds, and the excavation in general, without any assistance from the system developers. The archaeologist conducted two sessions, each with a different student, with both sessions completing without interrupt.

One student commented that she quickly became familiar with the excavation process at Monte Polizzo, even though she had not previously been introduced to the site. In contrast to her previous experiences, where she found it difficult to visualize and understand a dig site before actually arriving at the location, she felt that VITA gave her a fairly good understanding of the overall layout of the site and its most significant features. Both students expressed their desire to see a system like VITA used in their courses to get a full understanding of the 3D nature of dig sites. In addition, they felt that a substantial amount of data classification and analysis time was saved when using VITA, eliminating hours of sorting through various reports and books when attempting to obtain a deep understanding of a particular site. They stated that VITA proved to be especially good at providing such contextual information. The archaeologist stated that in current archaeological practice, little co-located collaboration (beyond regular meeting-style discussions) happens in analyzing excavation data, primarily because of the lack of adequate visualization and collaboration tools. However, she felt that in VITA, easy availability of data and rich visual representation of both spatial and temporal characteristics made the environment very conducive to collaborative exploration and discussion.

7. Conclusions and Future Work

We have presented VITA, a prototype mixed reality system, designed in cooperation with archaeologists, that allows them to collaboratively discuss and analyze a digital reconstruction of a dig site. VITA makes it possible for multiple users to walk around the virtual site, and explore it using multimodal interaction to inquire about interesting finds in situ. VITA also includes a collaborative table surface, augmented with a world in miniature model of the environment and high-resolution screens, to allow for simultaneous viewing of all available 2D and 3D site data.

Taking into consideration the limitations of the current cumbersome hardware, the preliminary user feedback has

been overwhelmingly positive and very enthusiastic. We are currently working with our archaeologist colleagues to incorporate the VITA system into the classroom curriculum and are planning to conduct a more thorough user study in the near future. As mentioned in Section 5, we also plan to explore ways to further personalize the user experience based on the user's level of expertise or interests, potentially allowing two users to visualize the same data set in different contexts.

Finally, we intend to explore how well our system can support remote collaboration. Since many excavations are a joint effort by multiple archaeologists from varying geographic locations, 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, and we need to determine how well it works in practice.

Acknowledgements

We wish to thank Trinity Jackman and Ian Morris and the Stanford Archaeology Center for their archaeological expertise, data, and access to the Monte Polizzo excavation site; Peter Allen, Alejandro Troccoli, and Ben Smith for the scanning data and 3D models; Lynn Meskell and James Conlon for their archaeological expertise; Sajid Sadi and Avinanindra Utukuri for support with the P5 glove; Lawrence Wang for his work on DiamondTouch code; 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 N00014-04-1-0005, N00014-99-1-0394, N00014-99-1-0683, and N00014-99-1-0249, and gifts from Mitsubishi Electric Research Labs, Alias Systems, and Microsoft Research.

Video

Work presented in this paper can be previewed in a digital video (DivX encoded) available for download from: www.cs.columbia.edu/graphics/projects/vita.

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," *Proc. IEEE Visualization 2001*, pp. 493–496, 2001.
- [2] Agrawala, M., Beers, A., Fröhlich, B., Hanrahan, P., McDowall, I., and Bolas, M., "The two-user responsive workbench: Support of collaboration through individual views of a shared space," *Computer Graphics (SIGGRAPH '97 Proceedings)*, pp. 327–332, 1997.

- [3] Allen, P., Feiner, S., Meskell, L., Ross, K., Troccoli, A., Smith, B., Benko, H., Ishak, E., and Conlon, J., "Digitally Modeling, Visualizing and Preserving Archaeological Sites [poster]," *To appear in Joint Conference on Digital Libraries 2004 (JCDL 2004)*, 2004.
- [4] "Archaeological Site Manual," 3rd ed. London: Museum of London Archaeology Service, 1994.
- [5] "ArcView GIS," ESRI, <http://www.esri.com>
- [6] "AutoCAD," Autodesk, <http://www.autodesk.com>
- [7] Baudisch, P., Good, N., and Stewart, P., "Focus Plus Context Screens: Combining Display Technology with Visualization Techniques," *Proc. ACM Symp. on User Interface Software and Technology (UIST '01)*, pp. 31–40, 2001.
- [8] Benko, H., Ishak, E., and Feiner, S., "Cross-Dimensional Gestural Interaction Techniques for Hybrid Immersive Environments," Columbia University, New York, Technical Report CUCS-029-04, August, 2004.
- [9] Bier, E. A., Stone, M. C., Pier, K., Buxton, W., and DeRose, T., "Toolglass and magic lenses: The see-through interface," *Proc. SIGGRAPH '93*, vol. 27, pp. 73–80, 1993.
- [10] Billinghamurst, 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, pp. 261–284.
- [11] Broll, W., Meier, E., and Schardt, T., "The Virtual Round Table - A collaborative augmented multi-user environment," *Proc. Collaborative Virtual Environments (CVE 2000)*, 2000.
- [12] Butz, A., Höllerer, T., Feiner, S., MacIntyre, B., and Beshers, C., "Enveloping Users and Computers in a Collaborative 3D Augmented Reality," *Proc. IWAR '99 (Int. Workshop on Augmented Reality)*, pp. 35–44, 1999.
- [13] Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A., "Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE," in *Computer Graphics (Proc. ACM SIGGRAPH '93), Annual Conference Series*, 1993, pp. 135–142.
- [14] Dietz, P. and Lehigh, D., "DiamondTouch: a Multi-User Touch Technology," *Proc. ACM Symp. on User Interface Software and Technology (UIST '01)*, pp. 219–226, 2001.
- [15] Feiner, S. and Shamash, A., "Hybrid User Interfaces: Breeding Virtually Bigger Interfaces for Physically Smaller Computers," *Proc. ACM Symp. on User Interface Software and Technology (UIST '91)*, pp. 9–17, 1991.
- [16] Gaitarzes, A., Christopoulos, D., and Roussou, M., "Reviving the past: Cultural Heritage meets virtual reality," *Proc. Conf. on Virtual Reality, Archaeology and Cultural Heritage*, pp. 103–110, 2001.
- [17] Harris, E. C., *Principles of Archaeological Stratigraphy*, 2nd ed. London, New York: Academic Press, 1989.
- [18] Hua, H., Brown, L. D., Gao, C., and Ahuja, N., "A new collaborative infrastructure: SCAPE," *Proc. IEEE Virtual Reality (VR'03)*, 2003.
- [19] "INTRASIS," The Swedish National Heritage Board, Archaeological Excavations Department and LandFocus IS AB, <http://www.raa.se/uv/intrasis/end/index.htm>
- [20] 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," *Proc. ICMI 2003 (Fifth Int. Conf. on Multimodal Interfaces)*, pp. 12–19, 2003.
- [21] Kitamura, Y., Konishi, T., Yamamoto, S., and Kishino, F., "Interactive stereoscopic display for three or more users," *Proc. SIGGRAPH '01*, pp. 231–240, 2001.
- [22] Krüger, W. and Fröhlich, B., "The responsive workbench," *IEEE Computer Graphics and Applications*, pp. 12–15, 1994.
- [23] Kumar, S., Cohen, P., and Levesque, H. J., "The Adaptive Agent Architecture: Achieving Fault-Tolerance Using Persistent Broker Teams," *Proc. Int. Conf. on Multi-Agent Systems*, pp. 159–166, 2000.
- [24] MacWilliams, A., Sandor, C., Wagner, M., Bauer, M., Klinker, G., and Brügge, B., "Herding Sheep: Live System Development for Distributed Augmented Reality," *Proc. IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR'03)*, 2003.
- [25] Olwal, A., Benko, H., and Feiner, S., "SenseShapes: Using statistical geometry for object selection in a multimodal augmented reality system," *Proc. IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR 2003)*, pp. 300–301, 2003.
- [26] "P5 Glove," Essential Reality Inc., <http://www.essentialreality.com>
- [27] Pierce, J. S., Conway, M., Dantzich, M. v., and Robertson, G., "Toolspaces and Glances: Storing, Accessing, And Retrieving Objects In 3D Desktop Applications," in *Proc. Symposium on Interactive 3D Graphics*, S. N. Spencer, Ed. New York: ACM Press, 1999, pp. 163–168.
- [28] 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," *Proc. SIGGRAPH '98*, pp. 179–188, 1998.
- [29] Rekimoto, J., "Transvision: A Hand-held Augmented Reality System for Collaborative Design," *Proc. Virtual Systems and Multimedia (VSMM '96)*, pp. 31–39, 1996.
- [30] "The Stanford University Excavations at Monte Polizzo, Sicily," Stanford Archaeology Center, <http://archaeology.stanford.edu/MountPolizzo/project.htm>
- [31] Stoakley, R., Conway, M., and Pausch, R., "Virtual Reality on a WIM: Interactive Worlds in Miniature," in *Proc. CHI '95*. Denver, CO: ACM Press, 1995, pp. 265–272.
- [32] Szalavari, Z., Schmalstieg, D., Fuhrmann, A., and Gerwautz, M., "Studierstube: An Environment for Collaboration in Augmented Reality," *Virtual Reality*, vol. 3, pp. 37–48, 1998.