Use of Live Video Overlay on 3D Data for Distributed Collaborative Review

by
Austin Seungmin Lee

B.F.A. Kookmin University 2007
M.F.A. Media Design, Art Center College of Design 2010

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

Master of Science in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

Signature of Author ______________________
Program in Media Arts and Sciences
May 6, 2013

Certified by ______________________
Hiroshi Ishii
Jerome B. Weisner Professor of Media Arts and Sciences
Program in Media Arts and Sciences
Thesis Supervisor

Accepted by ______________________
Prof. Patricia Maes
Associate Academic Head
Program in Media Arts and Sciences

Use of Live Video Overlay on 3D Data for Distributed Collaborative Review

Austin Seungmin Lee

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on May 10, 2013, in partial fulfillment of the
requirements for the degree of
Master of Science in Media Arts and Sciences

ABSTRACT

Using everyday skills, such as pointing and drawing freehand sketches, facilitates effective communication when reviewing visual information. However, for sharing three-dimensional (3D) data, it is difficult to support such approaches of using hands and real ink in a seamless way. This thesis proposes a new system design called AnnoScape as an approach to performing a remote collaborative review of 3D digital data using a live video overlay of the desktop image on the viewports of the 3D scene. The system's virtual viewports are controlled with tangible handles and can be left spatially in the 3D data space. The viewports can be shared with remote collaborators both asynchronously and in real time. The system allows distributed users to navigate shared 3D space individually or jointly (synchronizing the viewport); generate an overlay of the live video of hand drawings, physical objects, and printed images from the desktop surface with the viewport; and control the legibility of the visual contents. This spatial video overlay technique in the 3D data space allows distributed users to share the live annotations over the synchronized viewports. We report the prototype design and initial experiments to explore AnnoScape's possibilities through the scenario of having remote collaborators review the exterior site and interior reconfiguration of an existing architectural setting.

Thesis supervisor: Hiroshi Ishii
Title: Jerome B. Weisner Professor of Media Arts and Sciences, Program in Media Arts and Sciences

Use of Live Video Overlay on 3D Data for Distributed Collaborative Review

Austin Seungmin Lee

The following people served as readers for this thesis:

Thesis Reader ______________________________

Chris Schmandt
Principal Research Scientist, MIT Media Lab

Thesis Reader ______________________________

Ken Perlin
Professor of Computer Science NYU Media Research Lab

ACKNOWLEDGMENTS

I want to thank my academic advisor, Hiroshi Ishii for his guidance and amazing support and for motivating me greatly with his endless passion for Tangible Media Group's vision-driven research. I will never forget the creative guidelines from Hiroshi and the days and nights of rigorous research discussions we spent coming up with the concept and system design of our platform. I thank my thesis readers for helping me to push this thesis toward realizing its full potential. Having Chris Schmandt as a reader was extremely meaningful, as he was the advisor of Kobayashi at the Media Lab, who previously co-invented the ClearBoard with Ishii before joining the lab. Chris was one of the domain experts who deeply understood the philosophy behind CSCW (computer-supported cooperative work) projects such as the ClearBoard and TeamWorkStation, which were a huge inspiration for our design. Chris was a great supporter, and I could not have finished the research without his critical input. Ken Perlin has helped me to build the project to the next level by sharing his amazing body of work and insights.

Chigira from NTT and my UROP, Kojo, helped me greatly in the implementation of the core AnnoScape platform for this research. Chigira, as the main developer of the platform, took the lead in the core software programming, and Kojo supported some of the key features, such as image legibility control and image processing. Thanks to my amazing classmate, Sam Luescher, and all of the TMG members, Sheng Kai (Tony) Tang, Daniel Leithinger, Sean, Ryuma, Xiao, Lining, Jinha, Alex, who gave invaluable feedback during the design/evaluation process. Anna Pereira especially helped conduct the preliminary user study. Mary, our TMG admin, has always been supportive. Lining, Xiao, Jifei, and Ryuma from our group, as well as Gershon and Jinjoo from the lab, especially helped me to polish my presentation during my "critday." Thanks to Kshitij Marwah, Misha Sra, Valentin Heun, Anette Kapri, David Hayden, and many wonderful people I met at the Media Lab for being such great friends as well as some of the brightest and most talented peers I have met. Thanks also to Sophia Chang, Gonglue Jiang, Wikang Ryu, Jiwon, and Hyun, who provided advisements regarding collaborative architectural/3D design.

All of my Art Center advisors, including Anne Burdick, Phil Van Allen, Ben Hooker, and friends from Pasadena, such as Yuseung Kim, Hoon Oh, Hyemi Kim, Haemi Yoon, Yookyoung, Haelim, Haejin, Dee, Ina, Mikey T., Ping, Alex Cho, Hunter, Chris, and Yuin have changed me greatly as a designer and a researcher in many good ways. Peter Cho and Elise Co, who fully supported me, will always remain as my mentors in my heart. Further thanks go to Jiwon Park, Dookyoung, Jaewon, Hyun Chung and Kyoungjoo for being here in the United States as my closest buddies along with Jaeho, Woohee, and all of my other Kookmin University friends from South Korea. I feel lucky to have met my mentors from Seoul, Jaehyouk Sung and Minha Yang, who have taught me the fundamentals of graphic/interaction design.

Thank you to my mom and dad for their support and love. Also, thanks to my brothers, Seunghyun and Seungjae, for being my best friends. All of my other family members additionally deserve thanks for their prayers and support. Thanks Hongji! Finally, thank you Jesus and God for making all of this possible.

CONTENTS

LIST OF FIGURES

1 INTRODUCTION

1.1 MOTIVATION

Many people use remote collaboration systems, such as screen/desktop sharing programs (e.g. Timbuktu, join.me, Skype), or groupware (e.g. Revit, Digital Project, Google Docs) to review or collaborate on projects from various locations. However, in most conventional systems for sharing information over a distance, the virtual telepresence space is discontinuous to the physical workspace, and users often cannot take advantage of everyday skills, such as freehand sketching and pointing gestures. This constraint applies to both 2D and 3D information sharing in most conventional remote collaboration systems. There have been several research projects that explored shared drawing tools for supporting seamless integration of physical individual workspace into the collaborative virtual workspace (Ishii, and Kobayashi, 1992; Ishii, 1990). Yet the tools for merging the work flow between physical and virtual workspace for sharing 3D data has been considered relatively far less than the collaborative 2D platforms. For collaborators, the ability to have discussions while utilizing 3D virtual data (3D computer aided design [CAD] model or 3D map data) as well as two-dimensional (2D) content (photos or 2D maps) or moving images on a seamless platform has huge potential for creative collaborations. In the fields of urban planning, interior design, and medical practices, conversation about virtual 3D data is inevitable. Additionally, for electronic commerce or digital navigation maps, 3D information is helpful for the users to precisely understand the actual information.

In his seminal book, Robert McKim, describes how visual thinking is crucial in the diverse disciplines, such as design, architecture, visual arts, and science, and emphasizes how drawing, seeing, and imagining is an important cycle for the process of visual thinking (McKim, 1973). In this thesis, we aim to introduce a system design that enables designers, engineers, architects, and artists to share information in a collaborative virtual 3D workspace to support creative review process not limited to 2D contents. To motivate this, we started by looking at the relationship between visual thinking and the general design process for architectural practice. Based on interviews with five designers (three architects and two designers of exhibit space), we learned a common procedure included site research, collecting data, creating visual diagrams, ideation, design execution, design review, design iteration, and design refinement. Among these steps, the process of ideation and review require most sketching and collaborative visual explorations (Fig. 1).

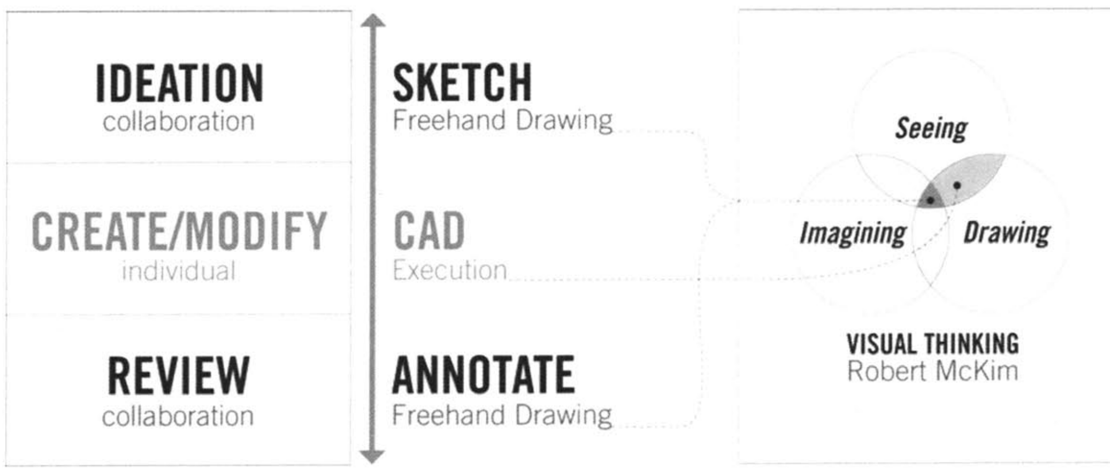

Figure 1: Design process, usage of sketching and visual thinking

The interaction during the design process can be largely categorized as below.
1) Ideation: In this stage, participants collaboratively collect data related to the problem, devise diagrams, and sketch to generate a collection of visual material that later could become the foundational elements for the actual design execution. The design practice using rapid freehand sketching can facilitate the process of idea generation and visual thinking.

2) Creation and Modification: Creating a 3D model from scratch requires advanced CAD skills (e.g. creating a plane, drawing a 2D virtual line on a virtual plane, and extruding or lathing). Modification of the 3D data includes moving and manipulating the existing model (e.g. grabbing and moving a building model to different places inside of a virtual 3D urban planning map). The creation and modification of the 3D content are mainly performed individually using state-of-the-art CAD software, which allows collaborators to work separately on different parts but seldom work on the same portion of the modeling simultaneously.

3) Review: The review stage includes interactions, such as shifting viewpoints, pointing (e.g. via physical finger pointing on the screen or using the mouse pointer), and leaving annotations on a 3D virtual scene (e.g. leaving quick annotations/comments digitally or physically on digital snapshots or paper plots from the 3D model). Rich collaborative interactions occur for the most part during the design practice (e.g. freehand sketching) and review phase of the 3D information (Moum, 2010).

While all of these interactions are important, we explicitly focus on improving the remote collaborative review experience of the 3D information and the sharing methods for the archived ideation sketches. In remote 3D collaboration, there has been a demand for a reviewing tool for discussion and commenting, versioning around the information. However, remote collaboration on a 3D project is challenging due to the ambiguity that can occur when sharing the 3D information. For example, from remote locations, it is very difficult to point to a model and to instantly annotate on a particular 3D scene in one workflow. Currently, none of the commercially available remote collaborative design review tools support shared freehand annotation techniques on the virtual 3D information that integrates the physical workspace into the virtual workspace. Our work explores interaction techniques, such as video overlay and viewport capturing in the virtual 3D workspace, to create a seamless navigation platform for collaborative review and presentation of 3D data over a distance.

1.2 OVERVIEW

Real-time sharing of annotations on 3D digital information, such as architectural /landscape models, has a large potential for architects and designers who are reviewing information over a distance. In conventional architectural practice for 3D collaboration, collaborators typically work with a mix of platforms (Fig. 2). For example, a common method for the remote review of 3D information includes a 2D print of the 3D scene for quick annotations or freehand sketching (Moum, 2010). In these conventional settings, discontinuity exists between the work platforms. There are several shared drawing systems that have shown effectiveness in displaying the physical workspace seamlessly in the digital workspace for remote collaborations, such as fusing paper media and digital documents together (Poupyrev, Tomokazu, and Weghorst, 1998; Ishii, and Miyake, 1991). However, workspace overlay techniques have been explored primarily to share and overlay 2D information. Our work specifically focuses on allowing users to integrate the individual physical desktop workspaces into the virtual shared workspace for reviewing 3D information over a distance.

Figure 2: An architect's desk. Photo courtesy of Anita Moum, author of "Design team stories: Exploring interdisciplinary use of 3D object models in practice."

We present AnnoScape, a remote collaboration system that allows geographically distributed users to navigate in the shared virtual 3D scene and to overlay a physical individual desktop workspace image to support individual and collaborative review of the 3D information by switching the work modes seamlessly (Figs. 5, 6, and 7). The interface consists of 1) Inter-Personal Space (IPS) for face-to-face conversations (Ishii, and Kobayashi, 1992; Ishii, and Miyake, 1991; Forlines et al., 2006) a 2) focused viewport that can merge into a synced workspace; and 3) a reference view for navigation, which shows the location of the focused viewports in the 3D environment from a bird's-eye view (Figs. 3 and 11). The hardware configuration includes a vertical screen, two webcams—one for capturing the participant's face and the other for capturing the physical individual desktop workspace from the top view, and a tangible camera controller used for changing the 3D scene (Fig.4, 9 and 10). Our main contribution is the system design that supports video overlay techniques in virtual 3D environments to allow distributed users to share the live annotations over the synchronized viewports for the collaborative review (Fig. 8). We report a preliminary user study on the usage of physical artifacts and freehand annotations during the collaborative review of 3D information over a distance, and we discuss the future direction of 3D remote collaboration.

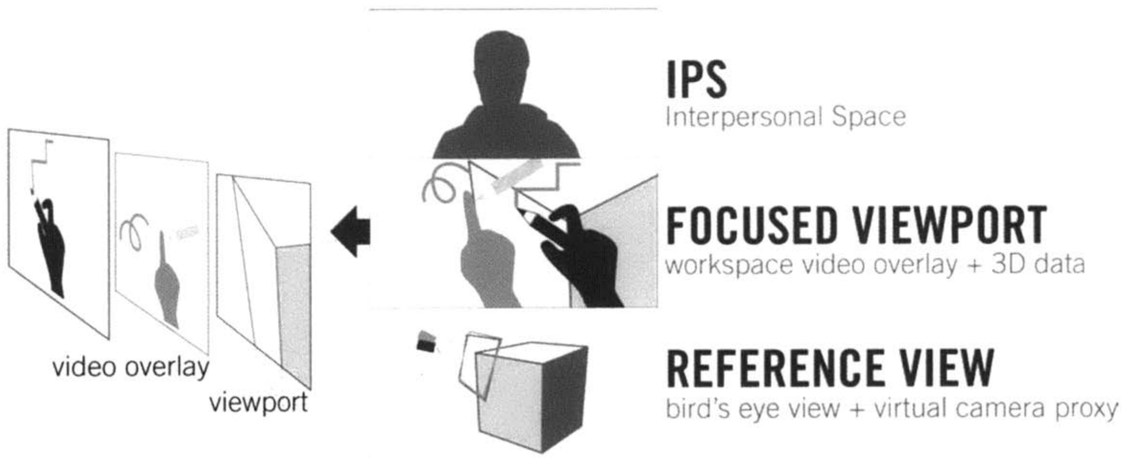

Figure 3: AnnoScape interface configuration (tightly coupled collaboration mode)

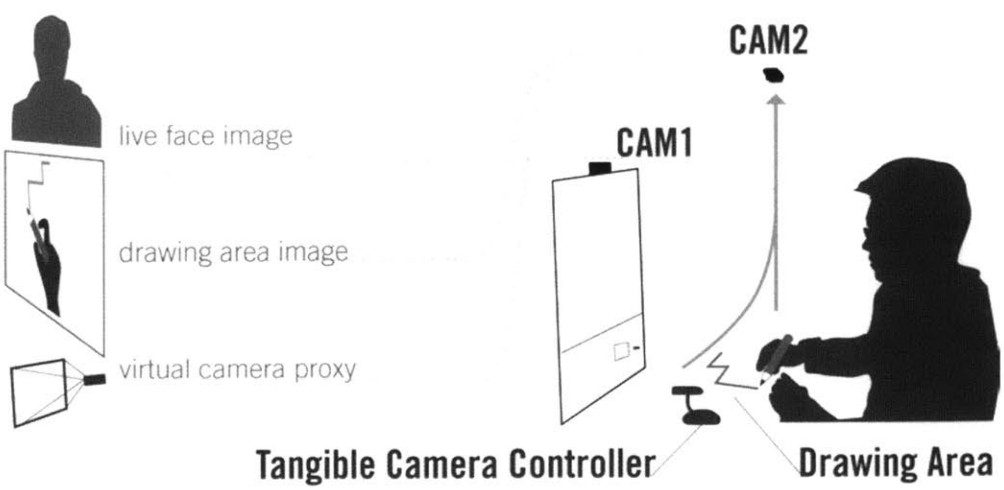

Figure 4: AnnoScape station hardware configuration

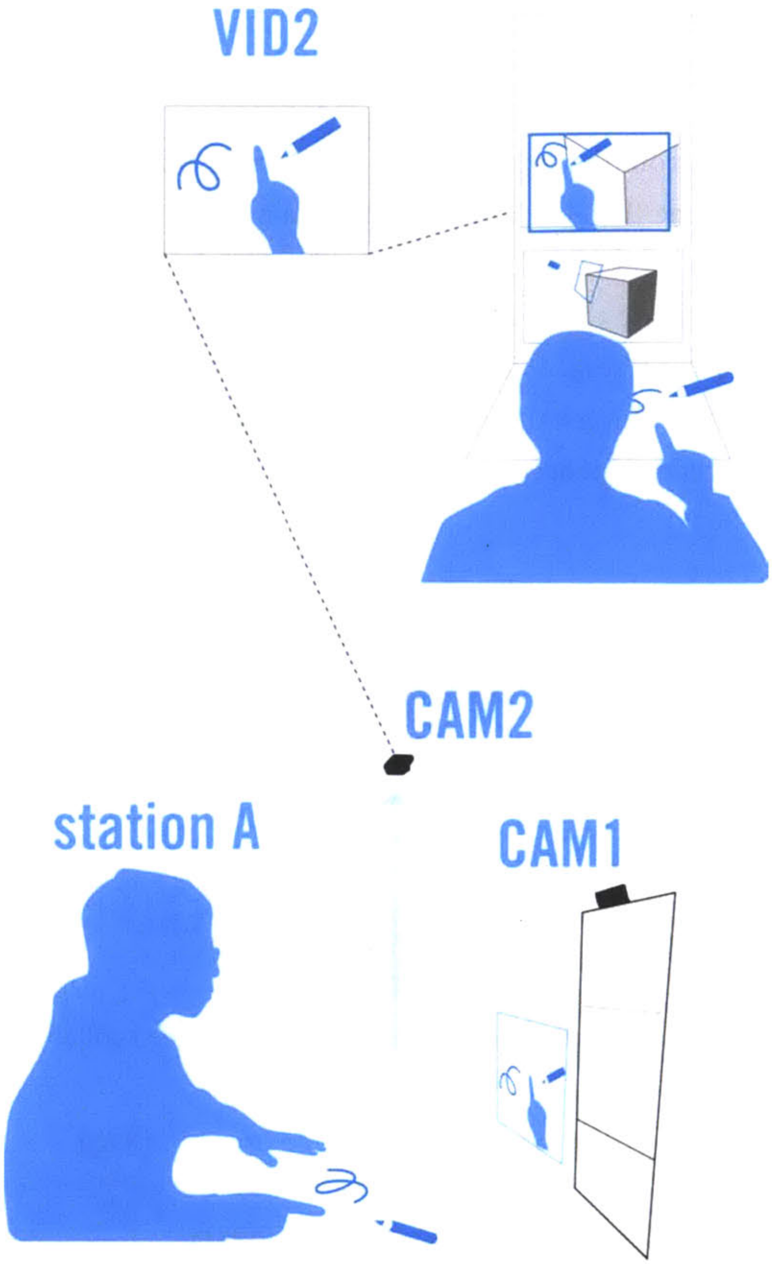

Figure 5: AnnoScape system infrastructure (individual review mode)

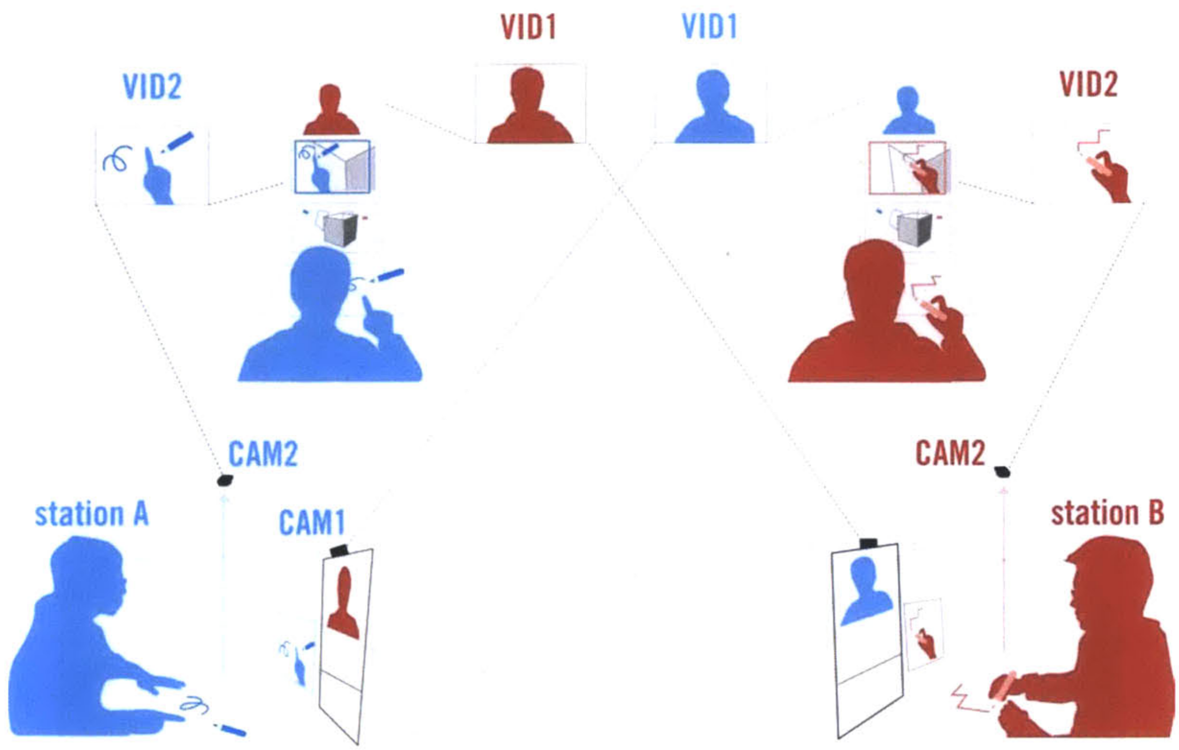

Figure 6: AnnoScape system infrastructure (loosely coupled collaboration mode)

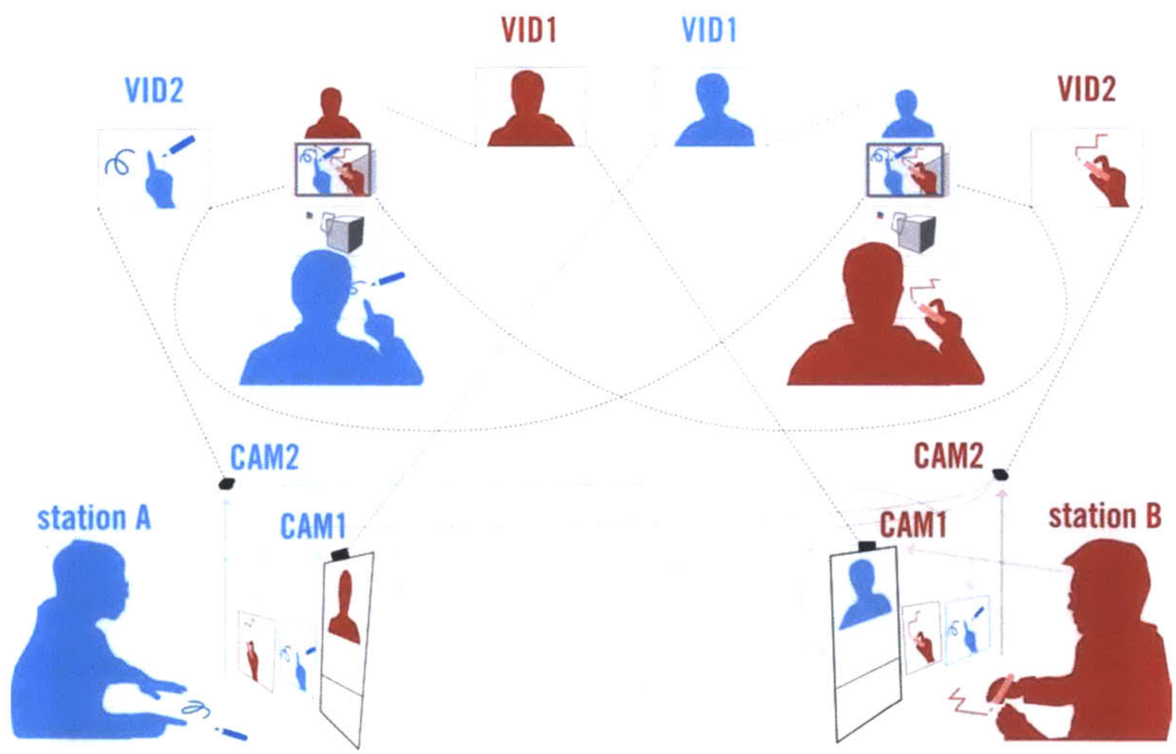

Figure 7: AnnoScape system infrastructure (tightly coupled collaboration mode)

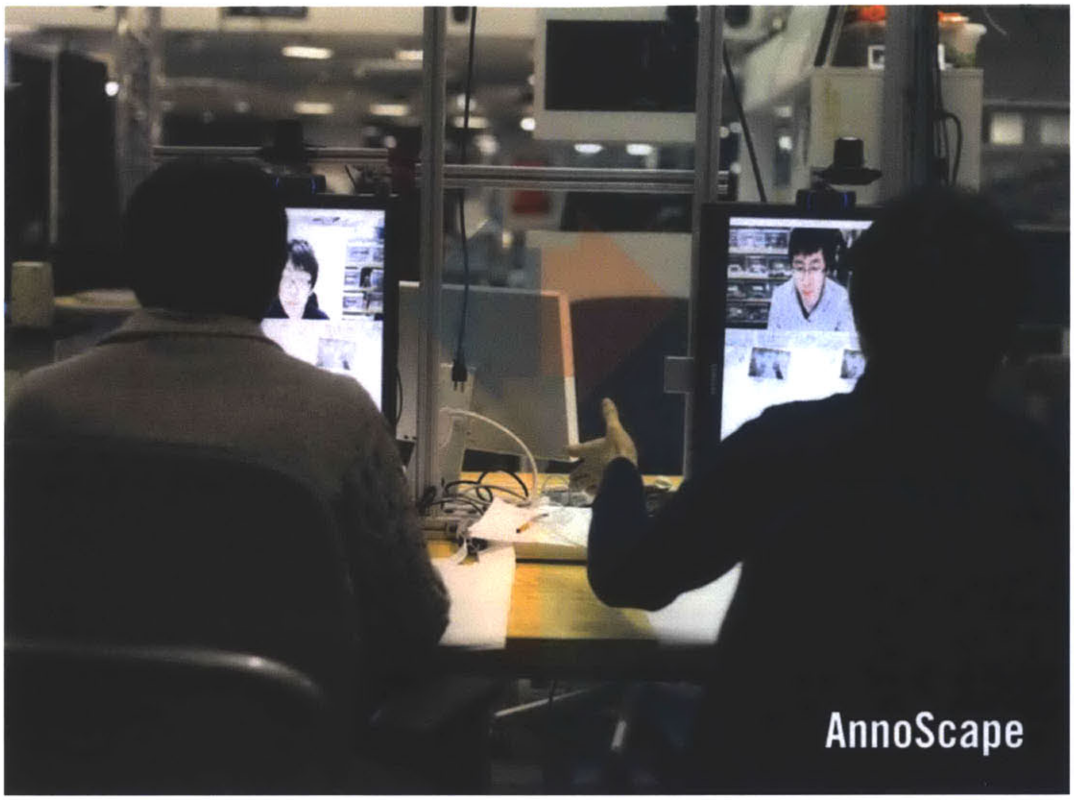

Figure 8: AnnoScape system

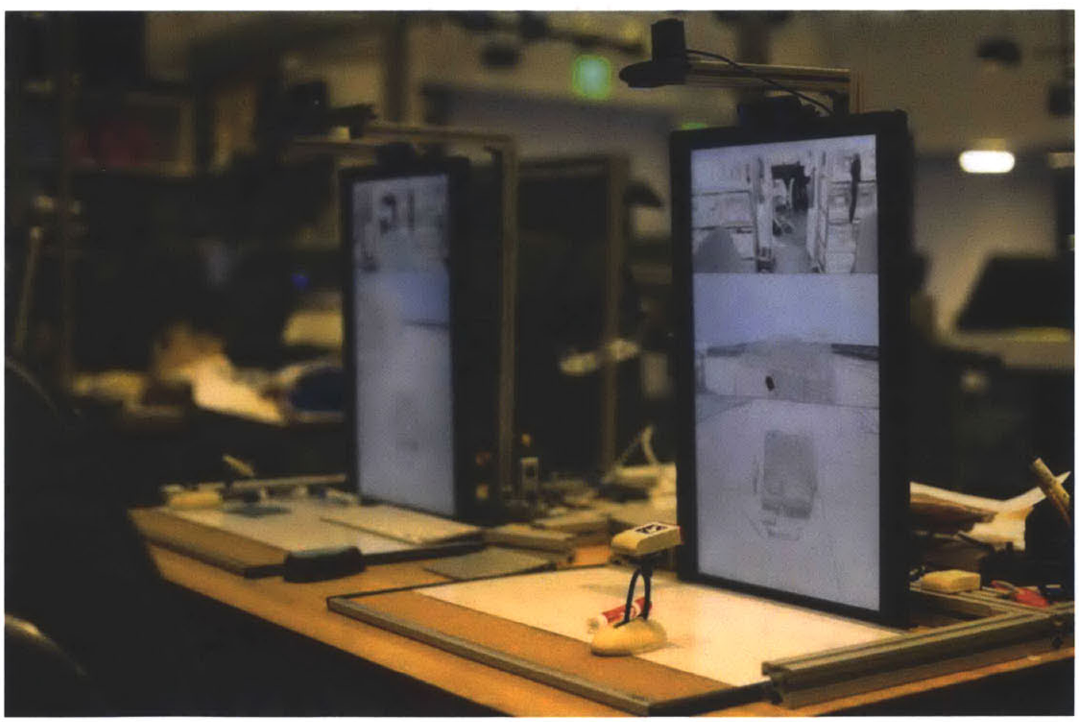

Figure 9: AnnoScape prototype setup

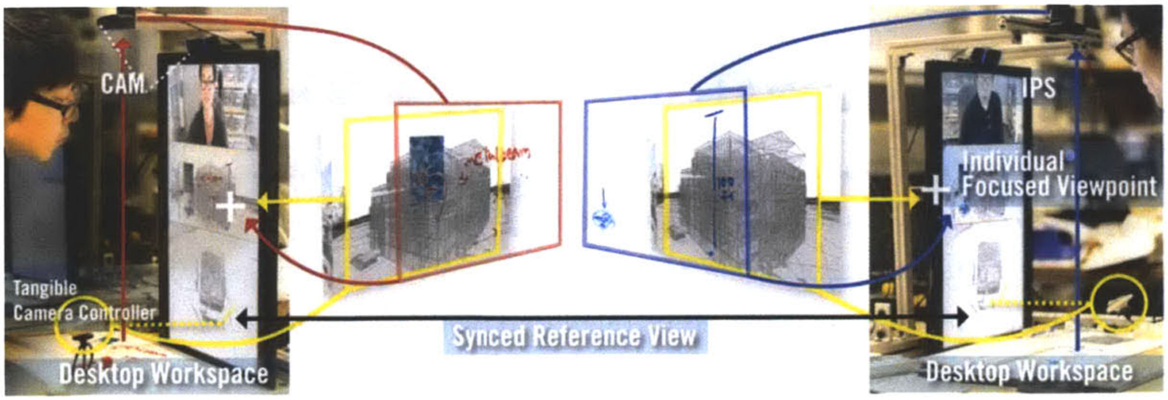

Figure 10: Loosely coupled real-time collaboration mode.

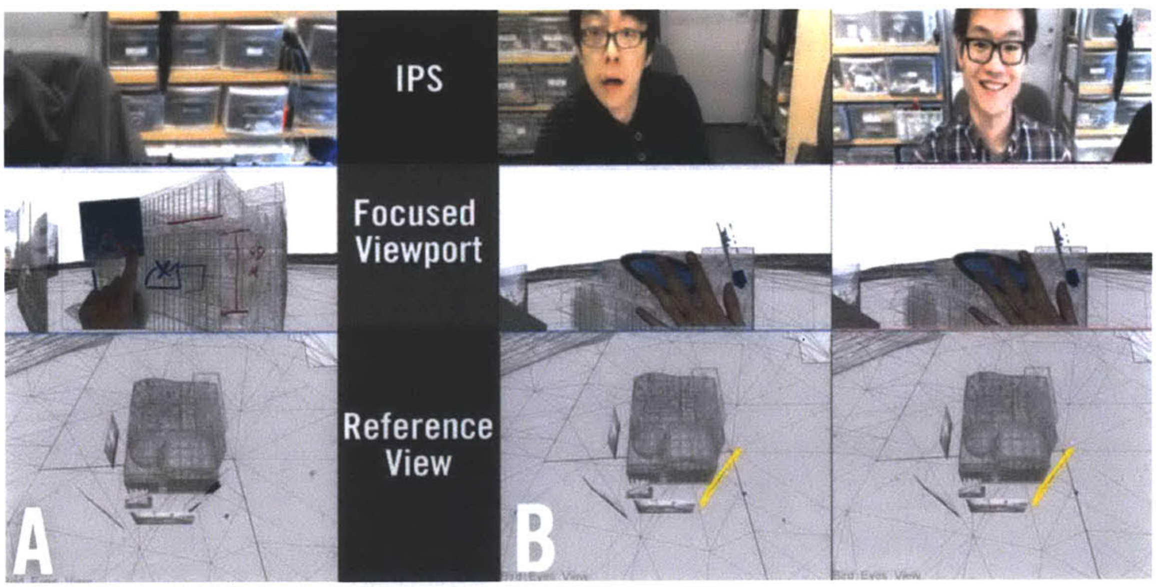

Figure 11: A) Individual Review mode, B) Tightly Coupled Real-time Collaboration mode

2 RELATED WORK

2.1 CONVENTIONAL 3D REVIEW

Three-dimensional information can be a virtual 3D model, physical 3D objects, and spatial coordinates in both virtual and physical space. In our system, we mainly focus on remotely sharing 3D information in virtual settings. Viewing the virtual 3D content can be achieved through two types of interactions: One is shifting the viewpoint and the other is manipulating the 3D object to see it from a different angle. However, to maintain the fixed position within the grid system for future 3D modeling and modifications, it is uncommon to manipulate the virtual object sorely for the purpose of viewing.

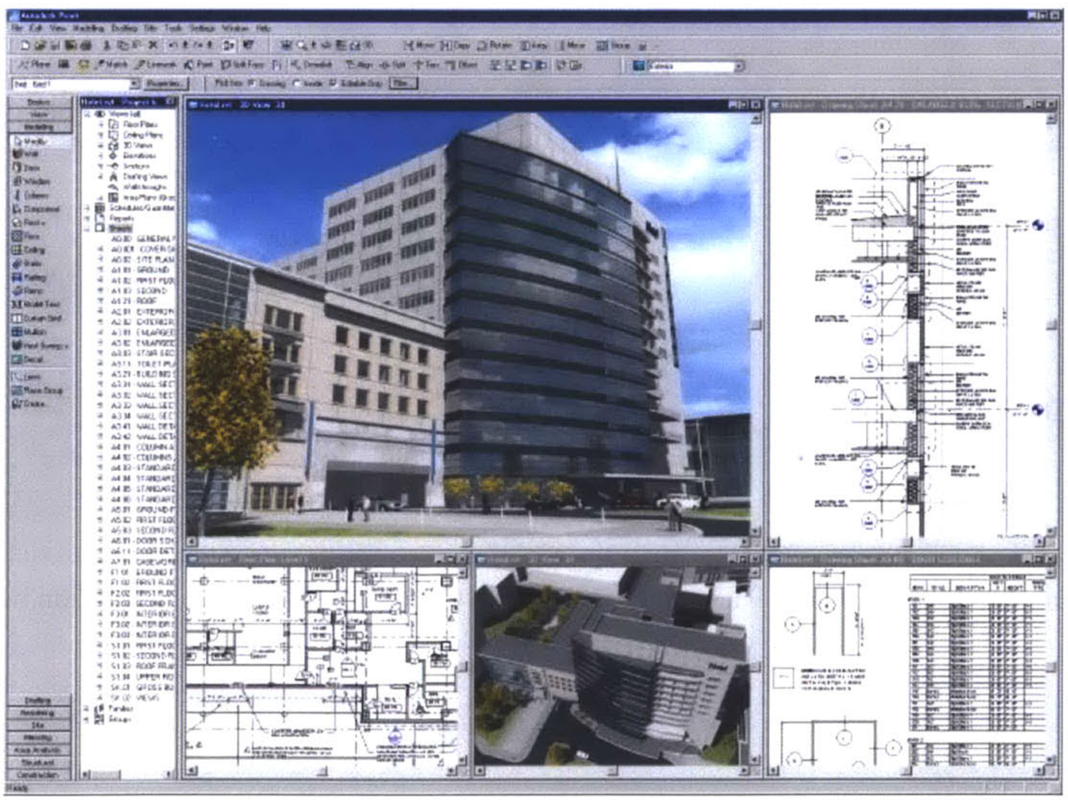

Figure 12: Revit is a high-end 3D groupware for collaborative digital modeling used in conventional architectural practice.

Currently, 3D viewing interaction examples can be found in many of the cutting-edge 3D CAD groupware programs, such as Autodesk's Revit (Fig. 12) or Gehry Technology's Digital Project, which enable professional architects to collaborate on a virtual model simultaneously (Revit; Digital Project) Although, these programs are powerful for executing the design for skillful CAD operators, they lack the key features used in many remote communications, such as face-to-face video support and freehand annotations. These limitations are compensated by additional video conferencing software, such as Skype and screen capturing systems or through the process of sending and printing the 3D scene for detailed reviewing. Another common method is using screen share and remote desktop control software, such as join.me, to review and manipulate the 3D model directly in the CAD system environment: however, this method is relatively low-end—the setting is limited to a single view and requires turn-taking for controlling the viewpoint. Furthermore, the fragments of information, including talking heads, shared content, and physical workspace using multiple applications, cause discontinuity in the workflow. All of these remote 3D collaboration platforms, from the high-end to the low-end, rely on sophisticated CAD tools that require additional integration of telepresence communication features. For asynchronous remote review and presentation of the 3D data, many designers would typically compile and share the digital snapshots of the virtual 3D scene in portable document format (PDF) and send the original 3D file via email.

2.2 NAVIGATING THROUGH 3D INFORMATION

2.2.1 Reviewing 3D in a Virtual Environment

There have been many explorations of reviewing 3D information in a computer-based, simulated, digital environment. Terravision, a predecessor of Google Earth, allows users to navigate spatially in a virtual representation of the earth and to review integrated spatial information, such as videos or photo images, in the 3D scene (art+com, 1994). Art+Com's system allows users to navigate from a large-scale view of the earth to very detailed scenes of buildings and objects (Fig. 13). Terravision's navigation is based on virtual camera control using a 3D mouse. Microsoft's PhotoSynth allows users to utilize captured images of physical surroundings to generate the 3D model of the photos and point clouds of the captured objects in a virtual environment (http://photosynth.net/). Users can navigate around the spatially connected photos and see the 3D picture from various angles (Fig.14). In the AnnoScape system, we also allow users to review the 3D information integrated with associated contents, such as concurrent snapshots of the annotation session, video, and static images.

Figure 13: Terravision(1994), image courtesy of Art+Com

Figure 14: Photosynth, image courtesy of Microsoft

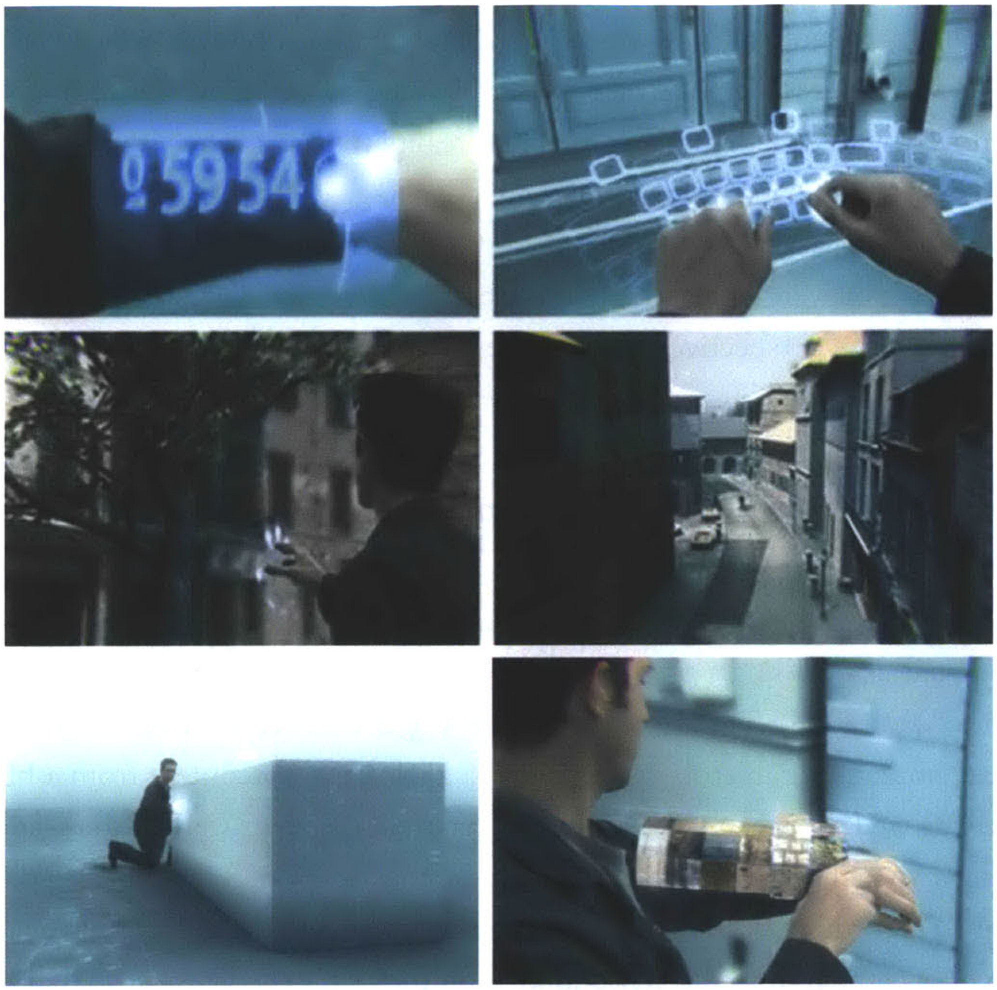

Figure 15: Bruce Branit's World Builder screen shots.

For collaborative 3D reviews over a distance, several approaches have used virtual avatars (Koutsabasis et al., 2012). Although the avatars help users to have a better sense of the scale of architectural structures, these systems introduce new types of complex interaction techniques that are irrelevant to the well-known interactions of conventional 3D collaborations. Collaborative Augmented Reality (AR) projects propose tangible interaction methods combined with AR display techniques (Billinghurst and Kato, 2002) . Although these systems allow users to take advantage of the physical space, the face-to-face communication issue remains difficult to resolve in many of the head-mount display applications, as the device typically covers the users' eyes. Ken Perlin's ARCADE (ARCADE) is a system for augmenting gesture-based computer graphic presentations. ARCADE's system allows the presenter to generate visual illusions of holographic 3D object interactions through real-time, video-based presentations. Bruce Branit's vision video, World Builder, demonstrates a user building holographic worlds in a virtual environment where the body is a simulated reality (Fig. 15). The vision of World Builder illustrates how a user can take full advantage of sensory awareness when reviewing and generating virtual 3D data. One possibility that World Builder did not explore, however, is the interaction techniques for reviewing the virtual 3D data from diverse perspectives, such as the bird's-eye view.

2.2.2 Vertical and Horizontal Display For 3D Review

Many research projects utilize the horizontal tabletop workspace for reviewing the 3D data (Ajaj et al., 2009). Forlines et al. introduced a multi-device, multi-user environment using a single-user commercial application for collocated users (Forlines et al., 2006). The work builds on Google Earth and includes a touch-sensitive, horizontal tabletop display for visualizing the 3D content from a bird's-eye-view as well as three-wall displays for multiple points of view of the scene. By dragging a virtual camera pointer through the touch-input on the tabletop display in X and Y coordinates, the user can control the scene on the vertical screen. The system allows the user to directly annotate on the bird's-eye-view scene that is displayed on the tabletop. The digital stroke annotation on the tabletop is geospatially registered in Google Earth and is displayed as a static transparent overlay in the accurate locations. Although the proposed application supports navigation and annotation in the 3D virtual space, the digital stroke overlay and the viewport navigation are limited to a single height due to the constraints of the touch-input interaction.

2.2.3 Reviewing 3D Using Tangible Controllers

Inspired by the concept of physical artifacts as input devices (Fitzmaurice, Ishii, and Buxton, 1995), BUILD-IT demonstrates the use of tangible handles to spatially manipulate a virtual camera in the 3D environment (Fjeld, Bichsel, and Rauterberg, 1998). The 3D scene is rendered on a vertical screen from the perspective view of this camera.

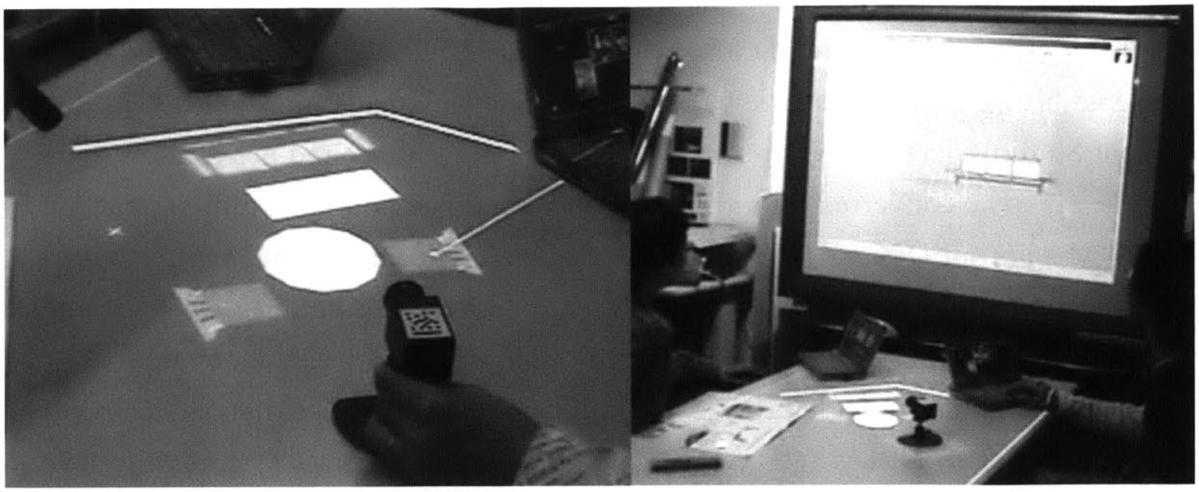

Figure 16: Rekimoto's Augmented Surfaces

In Augmented Surfaces (Rekimoto, 1999), a mock-up camera is used to review a virtual 3D scene that is projected on a tabletop (Fig. 16). The virtual scene is rendered on a vertical display based on the position and orientation of the mock-up camera. AnnoScape's 3D navigation is also based on the tangible manipulation of the virtual camera proxy inside the scene using physical objects placed on a tabletop workspace.

2.3 FREEHAND ANNOTATION IN 3D

We categorized 3D drawings or annotations by their relations with space. First is the spatial drawing, which includes forms such as helix or random curves in the 3D space. The second is the surface-driven annotations, which includes 2D drawings that can be spatially positioned in multiple angles in 3D. The applications, including spatial drawing, mainly focus on form-giving, whereas the latter application of 2D drawings in the 3D space often supports the design review process.

2.3.1 Spatial Drawings in 3D

Light drawing or camera painting is a well-known technique for generating spatial drawings. This photographic technique uses exposures made by a dynamic hand-held light source. Notably, Pablo Picasso applied the technique in art when Gjon Mili introduced his photos of ice skaters with lights attached to their skates (Pablo

Picasso, 1949). Although the input of light drawings and paintings is produced with movement of light in 3D, the result is often static 2D photos, which makes it impossible for viewers to review the generated spatial images in 3D (Fig. 17). Front Studio's Sketch furniture is a system that allows freehand pen strokes made in the air to be converted into 3D digital files through Motion Capture recording (MOMA | Front Design). The designers can later 3D-print the midair sketch into physical furniture (Fig. 18). While these spatial drawing techniques are well-suited for producing rapid prototypes or abstract forms in 3D, for beginners, without having direct visual and haptic feedback, it can be less accurate than sketching on a flat surface. 3Doodler is a 3D printing pen that allows users to generate physical forms in 3D space in real time. The system allows users to spatially draw arbitrary forms through freehand pen strokes. Although it is possible to create sophisticated 3D forms with expert crafting skills, it is complicated to leave annotations spatially in midair using the system.

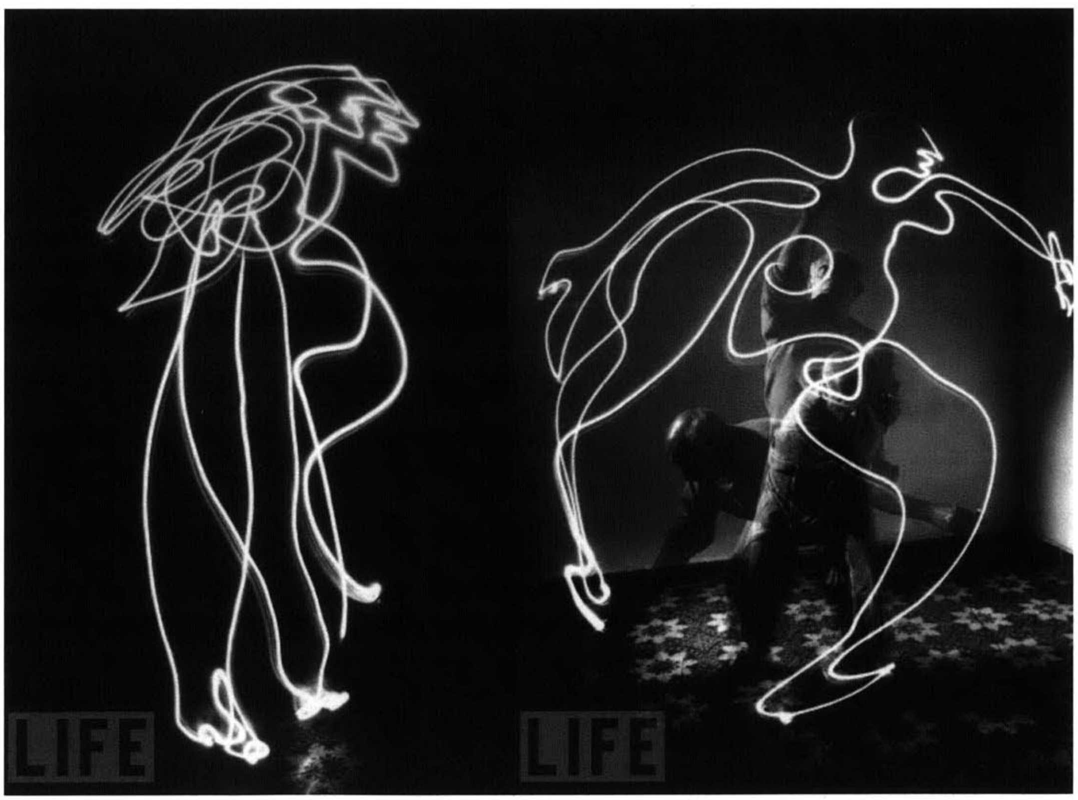

Figure 17: Picasso light drawing, photo courtesy of *LIFE* magazine

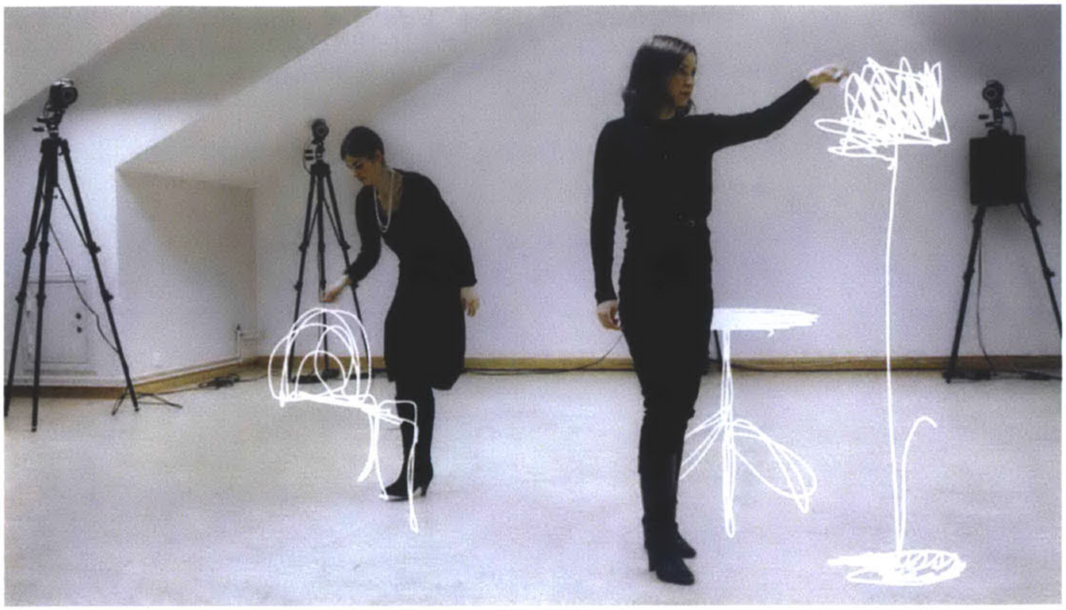

Figure 18: Front Studio's Furniture Sketch

2.3.2 Surface-Driven Annotation in 3D

There have been several approaches to create freehand annotation in the 3D virtual space in the research community. Virtual Note Pad (Poupyrev, Tomokazu, and Weghorst, 1998) showed the potential of annotating on top of a 3D virtual medical environment using a head-mounted display. Boom Chameleon by Tsang et al. (Tsang et al., 2002) uses a spatially aware display to define the 2D plane for annotation and to use the video playback feature (Fig. 19). Both of the proposed applications support single users as they navigate and annotate using digital strokes in the 3D virtual space. However, they do not combine physical information such as gesture-pointing or real ink with the digital data. Merl's System proposed a 3D map navigation system on Google Earth (Aliakseyeu et al., 2002). It also supported the annotation on top of the map, but its annotation feature is limited to only the horizontal plane. Other systems have demonstrated collaborative annotation in the 3D space for collocated participants using tablets (Kasahara, Heun, and Lee, 2012).

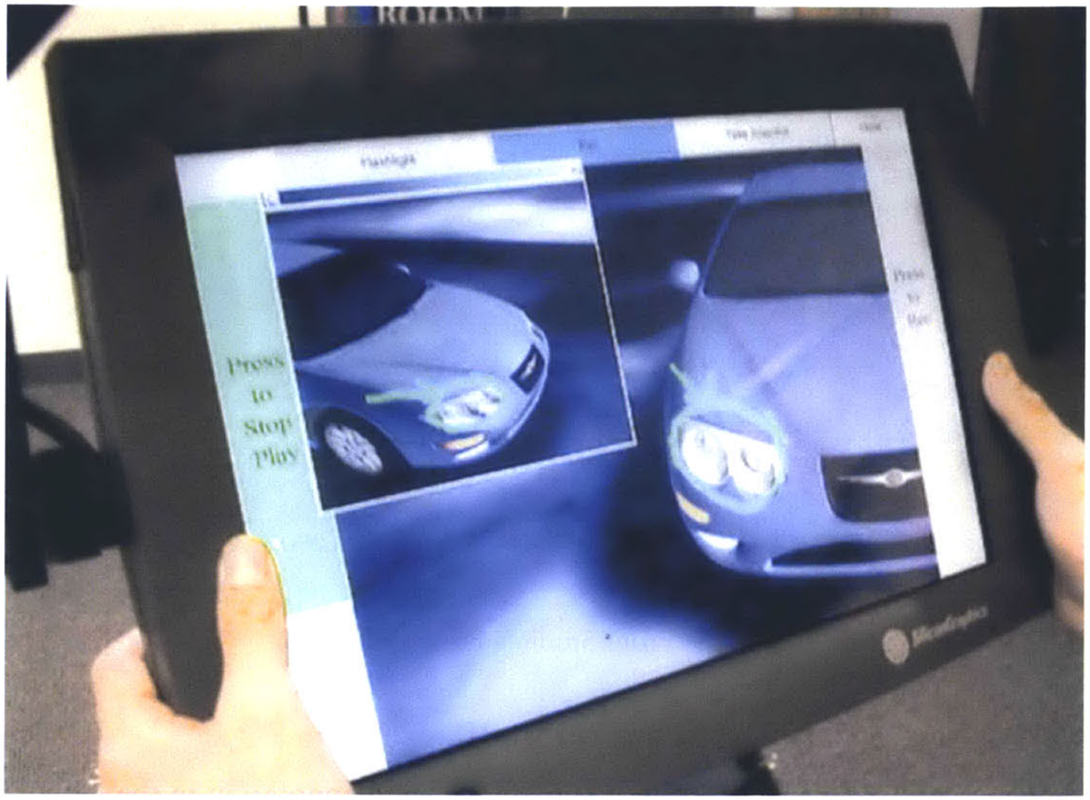

Figure 19: Boom Chameleon uses spatially aware display for navigation and annotations

Figure 20: Second Surface system allows users to collaboratively place drawings, texts, and photos three-dimensionally in physical locations using tablet-based AR technology.

T(ether) from Tangible Media Group at MIT Media Lab is a tablet-based application that allows users to draw directly on the surface and then move the tablet to spatially interact with volumetric data ("Tangible Media Group"). The system uses Vicon motion-capture cameras to track users' heads and hands so as to allow collocated collaborators to generate 3D contents using spatial drawing and to leave annotations in space. Second Surface is another tablet-based application that uses image-based AR recognition technology to enable users to generate annotations in the 3D space in everyday settings (Fig. 20).

In our current work, we concentrate on supporting the review process of collaborative 3D projects as opposed to generating 3D forms. To achieve this, our system design focuses on integrating 2D annotations in the 3D space. We combine the video overlay-based shared drawing techniques with virtual 3D review to achieve freehand annotation on the 3D data. A significant benefit of the spatial video overlay approach is that the system allows remote participants to share the increased content from the individual desktop workspace, including hand gestures and freehand annotations, in virtual 3D environments.

2.4 Video Overlay in Collaborative Reviews

2.4.1 Shared Surface Applications

The workspace overlay technique for remote collaboration has been explored primarily on remote 2D collaborations. Sharing the live video of the physical workspace is one possible way of achieving the remote annotation (Fig. 21). This approach to 2D video overlay techniques has been achieved in several research projects that include ClearBoard (Hiroshi Ishii and Kobayashi, 1992), TeamWorkStation (Ishii, 1990), Double Digital Desk (Wellner, 1993), and IllumiShare (Junuzovic et al., 2012).

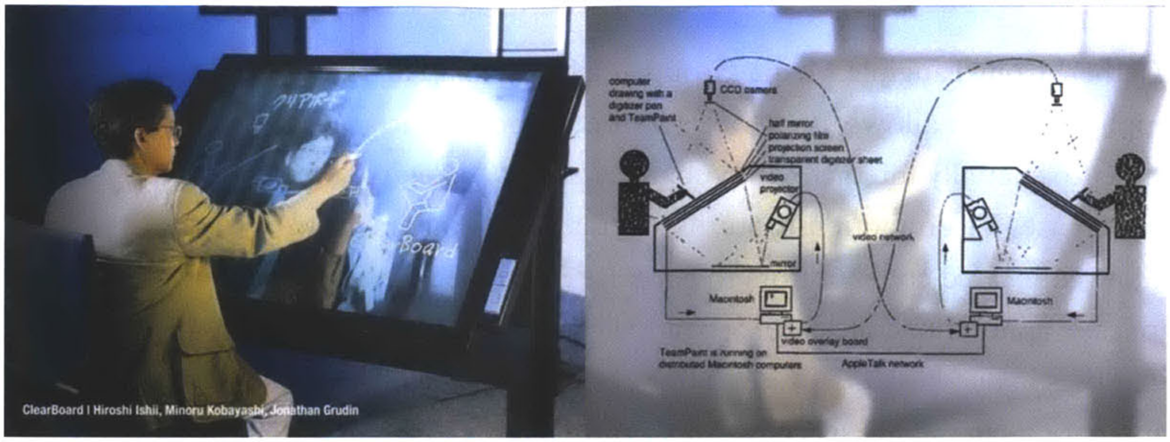

Figure 21: ClearBoard supports the gaze awareness feature and collaborative annotation in a seamless way.

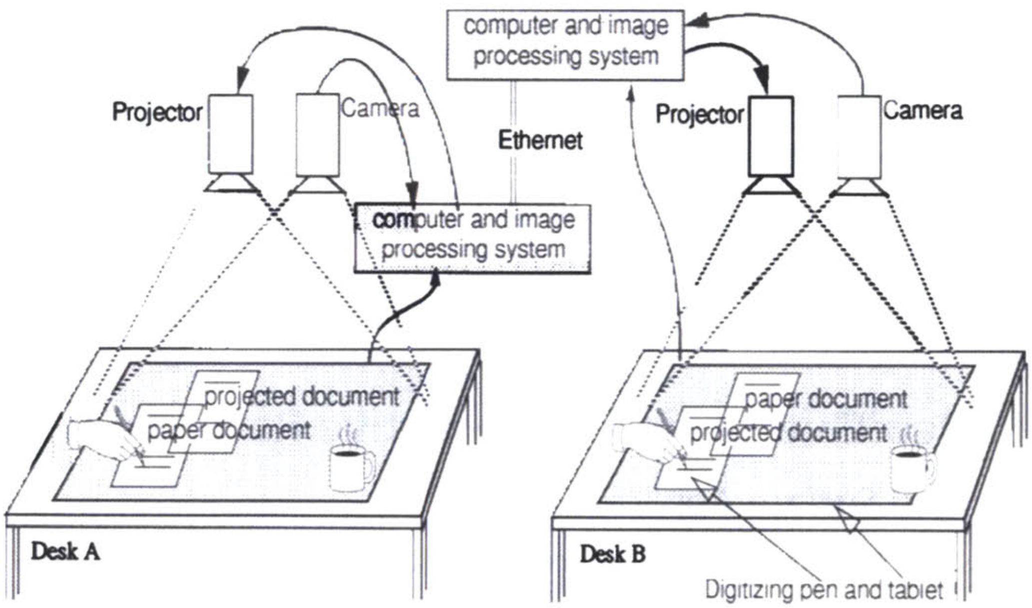

Figure 22: The Double Digital Desk

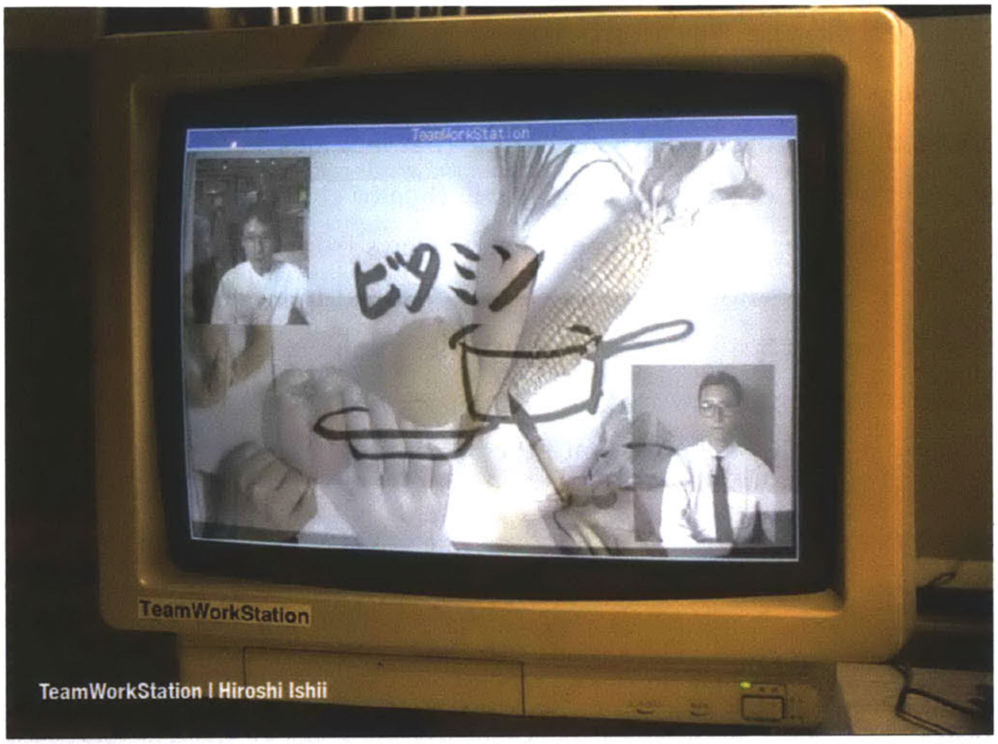

Figure 23: TeamWorkStation overlays individual desktop workspaces in the virtual workspace.

The TeamWorkStation's core idea is the "overlay of individual workspace image in a virtual shared workspace." The shared drawing surface of TeamWorkStation demonstrates variations in the overlay of individual workspace images, such as the screen overlay, the physical desk overlay, and a mix of screen and physical desk overlays (Fig. 23). Wellner's Double Digital Desk allows users to share paper documents over a distance as if they were sitting at the same workspace (Fig. 22). Similar to Double Digital Desk, Illumishare takes the approach of sharing arbitrary physical and digital surfaces by directly projecting the remote collaborator's shared surface on the local workspace. Pictionaire is an interactive tabletop system for collaborative design work for collocated users (Hartmann et al., 2010). The project uses overhead image capture and projection to generate digital 2D copies of physical 3D objects on its tabletop surface. Pictionaire demonstrates how annotations using physical objects can be effective for collaborative ideations. While these shared drawing systems provide a platform for quick expressions using annotation and gestural pointing on the 2D information, none of them are specifically designed for reviewing 3D contents.

2.4.2 Shared Space Applications

MirageTable provides 3D teleconferencing experience to its users using a depth camera and curved projection AR technique that includes 3D glasses, to capture and to share physical objects remotely in 3D (Benko et al., 2012). Although MirageTable provides correct 3D perspective views from the user's viewpoints (Fig. 24), the system is not designed with support features such as annotations or WISIWIS (What You See Is What I see) features (Stefik et al., 1986) for reviewing the 3D information. In our work, we focus specifically on the 2D annotation application that can be integrated into the 3D virtual space for remote reviews of existing 3D data.

Figure 24: Microsoft's Mirage Table

3 DESIGN PRINCIPLES

3.1 DESIGN OBJECTIVES

AnnoScape aims to provide a system that allows remote participants to collaboratively review 3D information and to leave 2D annotations spatially in the virtual 3D workspace. Our design principle is that AnnoScape allows users to maintain the individual usage of the traditional desktop workspace while reviewing the 3D data in the virtual scene. Our guiding principles were based on interviews on remote collaboration with five designers (three architects and two designers of exhibit space). Our guidelines are as follows:

1) Physical artifacts (e.g. pen, print, and sketch models) are essential during the design practice (e.g. freehand sketching, volume studies) and review phases of the information.

2) Collaborative review and annotation often require a mix of media such as the digital scene or paper plot from the 3D scene.

3) Users should be able to maintain their own work styles and tools for collaboration through the preservation of individual workspaces.

4) Users should be able to switch between different work modes with smooth flow.

5) Shared information should be legible and clear.

3.1.1 Interface Design Guidelines for Video Overlay

Tyler DeWitt is a science educator and an advocate of communicating science in simple, understandable way ("Tyler DeWitt: Hey Science Teachers -- Make It Fun | Video on TED.com"). He shares his science videos on YouTube to explain complex scientific information in a comprehensible way by capturing a desktop image with physical props and drawings that he points or moves around with his hands. Although the science education videos are made for unidirectional presentation, the configuration of Tyler's video interface is interestingly similar to that of TeamWorkStation (Fig. 25), which consists of IPS (inter-personal space) for capturing the live face images and the virtual shared workspace where the individual desktop images of the remote collaborators are overlaid together. The configurations that display hand gestures and live face images suggest how sharing live video images of the desktop workspace can help remote viewers to understand and communicate ideas effectively. Our hypothesis is that this technique can be applied in architectural collaborations and in other remote 3D project reviews since they usually include not only sharing the 3D CAD files on the screen but also working with paper plots or physical models placed on a physical desktop workspace. Currently, in conventional architectural collaborations over a distance, the workflow between the physical desktop and the screen is disconnected. Thus, the remote 3D review session is typically limited to sharing the viewpoints of the computer screen or to sending PDFs of digital snapshots of a 3D scene with annotations.

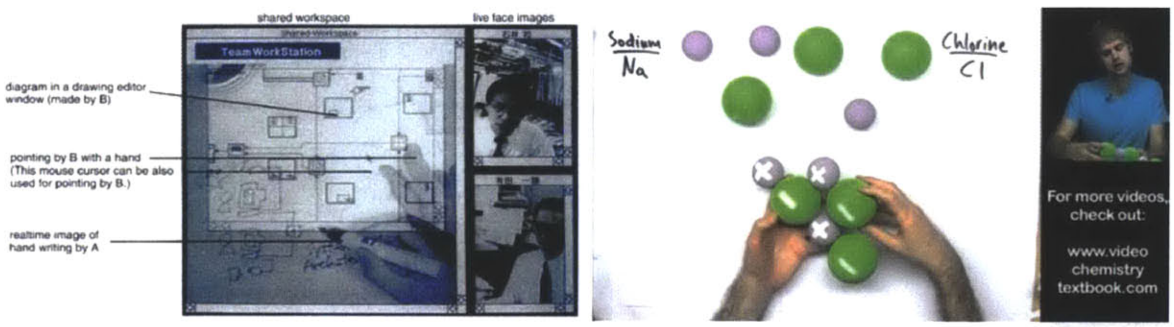

Figure 25: Comparison between TeamWorkStation interface (left) and Tyler's video (right)

3.1.2 Desktop Workspace Contents and Sketch Input

There have been many applications for annotating in the 3D space that utilize digital sketching processes, and some of the examples are covered previously in Section 2.3.2 on "Surface-Driven Annotation in 3D." Although digitized sketching provides clean and crisp strokes, at the same time, by eliminating the visual elements in the physical world, it loses the rich interactions that can be achieved in traditional desktop workspaces, such as finger pointing or moving physical materials during review sessions. Our system attempts to maintain the benefits of a traditional sketching process on the physical desktop workspace, which includes the tactility and familiarity provided by the basic desktop tools. As seen in TeamWorkStation (Fig. 20) and Pictionaire, the overlay images created by combining traditional drawing tools and physical objects can provide flexibility for rapid visual communication and ideations. However, in shared drawing applications based on the video overlay technique, there are issues such as legibility and identifying the owner of the captured images.

In the AnnoScape system, we aim to allow users to spatially overlay the live physical desktop image in the virtual viewport with high legibility by extracting the sketch or physical materials from the foreground through image processing. Also, the system provide tools for content-filtering options that can display captured images or associated contents with different alpha values to help users to identify the contents based on the owners or the time during which it was created. The live annotation sessions captured from the physcial desktop workspace can be viewed from various angles in the virtual 3D space.

3.2 DESIGN INTERACTIONS

3.2.1 Viewports and AnnoScape Interface

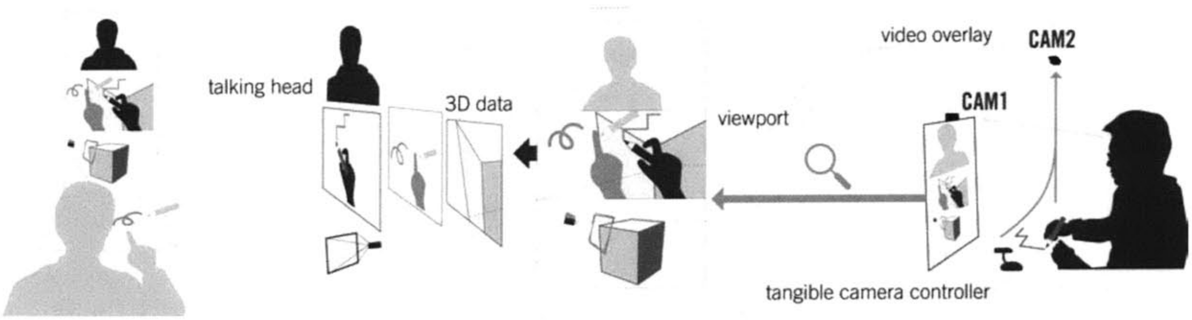

Figure 26: Viewports and AnnoScape interface flow diagram (tightly coupled)

In AnnoScape, the 3D scene is displayed in multiple viewports; the focused viewports for the workspace and the reference view that helps users to locate the focused viewport during navigation (Fig. 26). The reference view displays where the virtual camera proxy is located, and as the virtual camera moves, the focused viewport scene changes (Fig. 27). Each participant can navigate to a scene and capture the focused viewport location by pressing the capture button. In our system, capturing the personal focused viewport means capturing the scene as an area for annotation. In the captured mode, the focused viewport becomes the platform for the individual's virtual workspace. Once the scene is locked, the overhead web cam simultaneously captures the physical individual desktop workspace, and the system overlays the live video image onto the focused viewport. This allows the user to leave an annotation from the specific point of view on the captured 3D scene. Remote collaborators can join the annotation session by overlaying their focused viewport to create a synced workspace on the captured scene. As the participants leave the workspace, the annotation can be archived spatially in the virtual 3D space. In addition to the focused viewport and Reference View, our system interface also includes an IPS that displays the live face image of the collaborator.

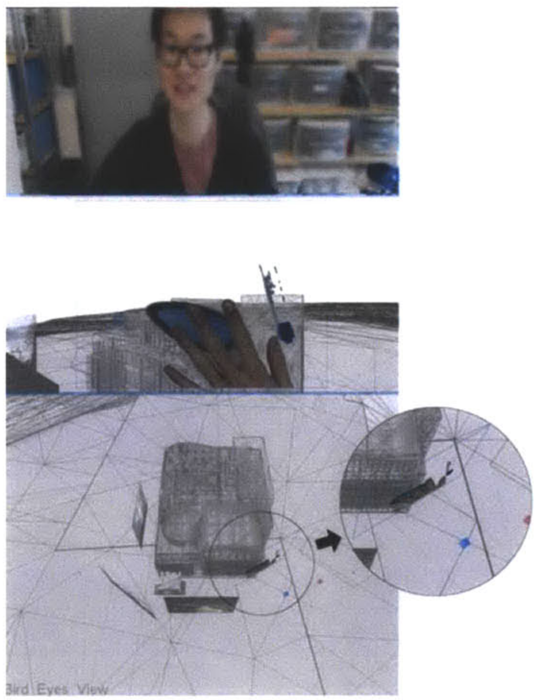

Figure 27: Live thumbnail image of the video overlay in the Reference View

Based on the work modes, the interface can be dynamically modified from 1) focused viewport + reference view for individual review mode (Fig. 28 and 29), to 2) IPS + focused viewport + reference view for loosely coupled real-time collaboration mode and default mode (Fig. 30 and 31), to 3) IPS + focused viewport for tightly coupled real-time collaboration mode (Fig. 32 and 33). For each type of work mode, our platform provides three selectable interfaces that provide five variations of work styles in our current prototype setup: a) The loosely coupled collaboration interface is the default interface, which consists of IPS + focused viewport + reference view; b) Custom individual review interface with focused viewport + reference view; and c) custom tightly coupled collaboration interface with IPS + focused viewport. The participants can switch between the default interface and the customized interfaces for each work mode.

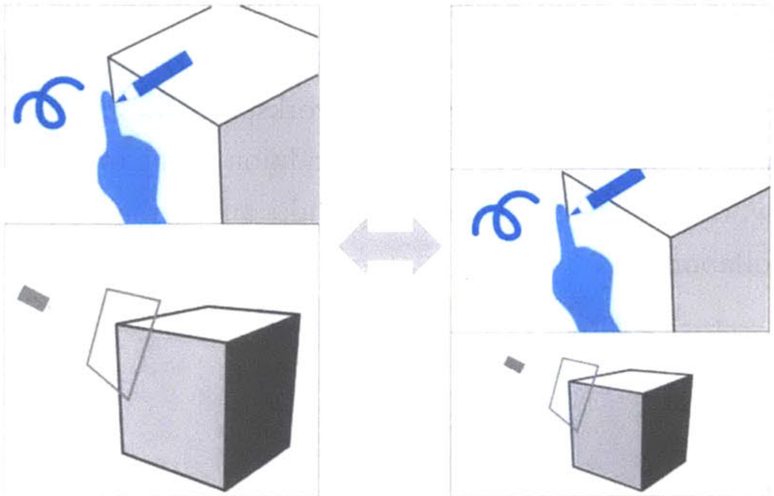

Figure 28: Individual mode, interface configuration options.

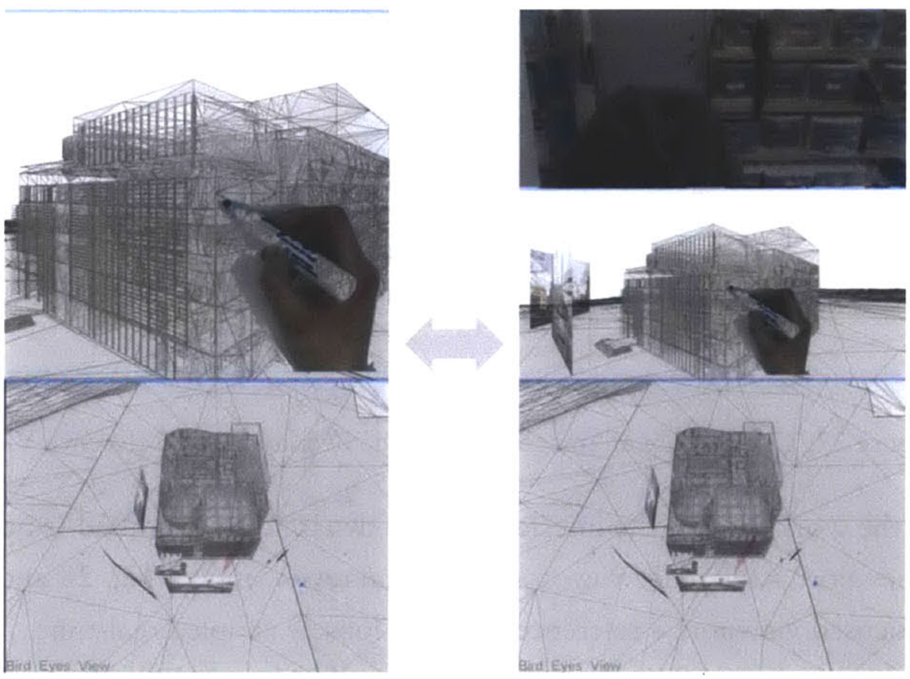

Figure 29: Custom individual review interface (left) and default interface (right): Individual review mode allows the user to turn off the IPS section by switching into the custom individual review mode, which divides the screen into two sections: the focused viewport and reference view (left).

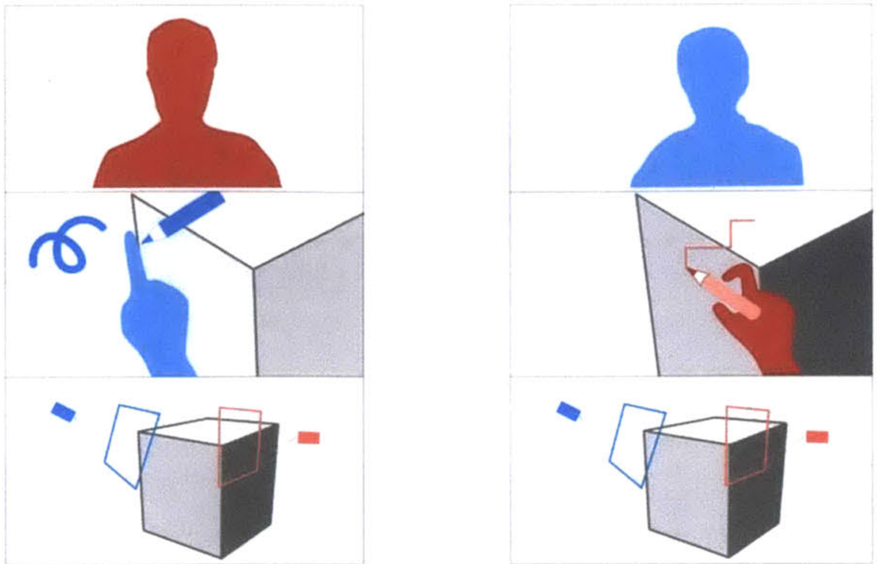

Figure 30: Loosely coupled real-time collaboration as the default interface configuration.

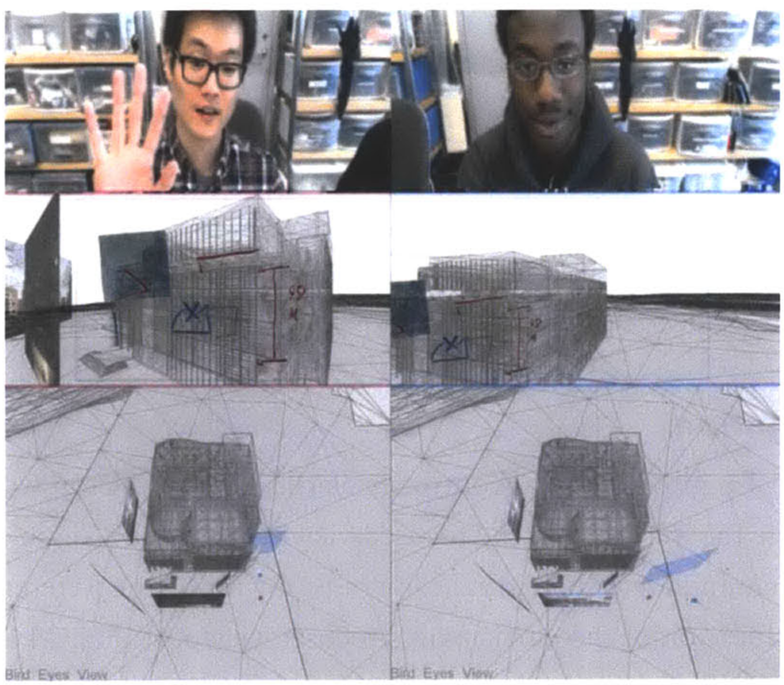

Figure 31: Loosely Coupled Collaboration (default) Interface: The loosely coupled real-time collaboration interface is the default mode, which consists of all three sections: IPS, focused viewport and reference view.

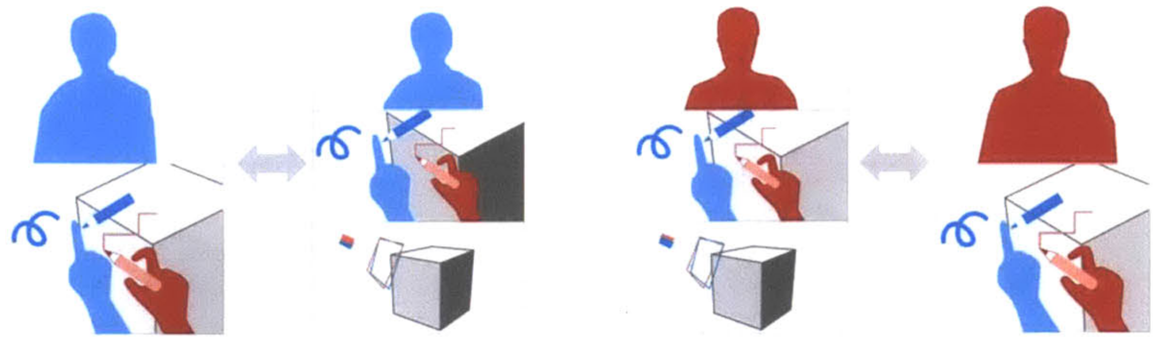

Figure 32: Tightly coupled real-time collaboration interface options.

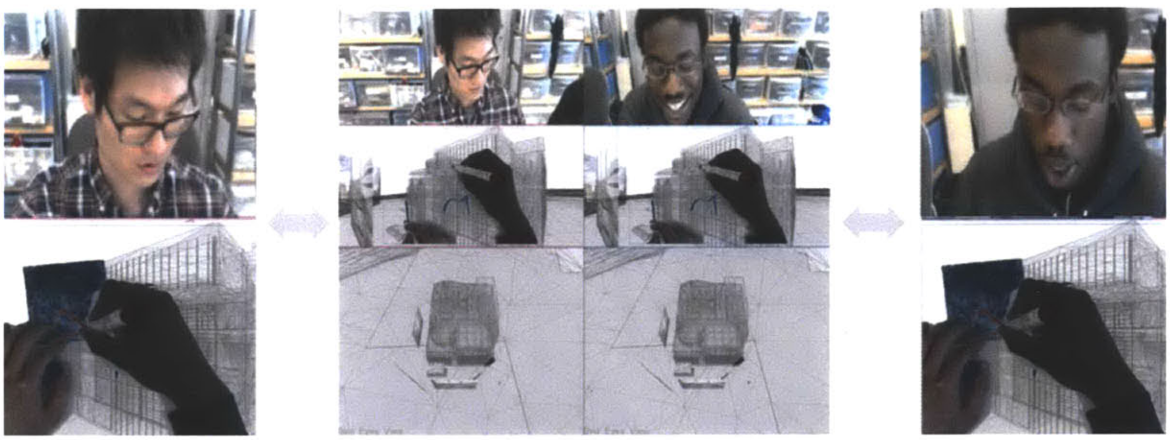

Figure 33: Tightly coupled collaboration interface and default interface (middle): Participant could switch the default interface to tightly coupled interface that displays IPS and scaled up focused viewport for better legibility.

3.2.2 Navigation

By moving the Tangible Camera Controller placed on the tabletop workspace, the user can change the 3D scene in the focused viewport. As the Tangible Camera Controller is manipulated, the virtual camera proxy inside the Reference View shows the 3D location of the participant's focused viewport (Fig. 34 and 35). When a remote collaborator starts the annotation, the virtual camera proxy spatially displays a thumbnail of the live streaming video showing the annotation session on the physical desktop. Participants can snap to the other collaborator's workspace by moving their personal viewport on top of the existing workspace.

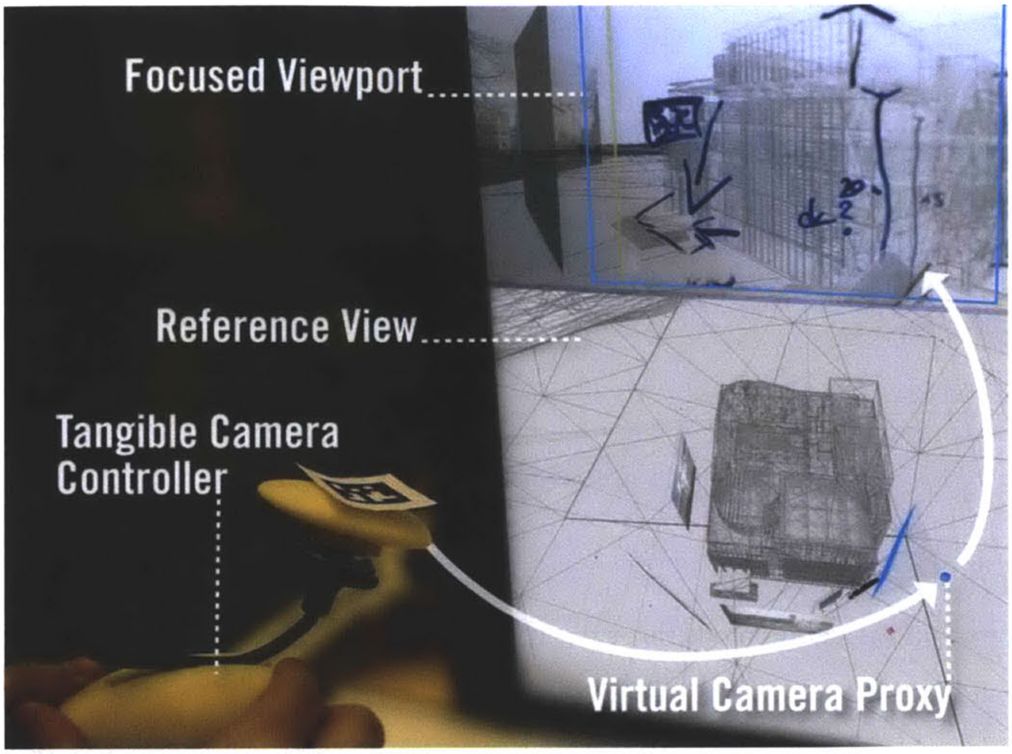

Figure 34: Tangible Camera Controller for navigation

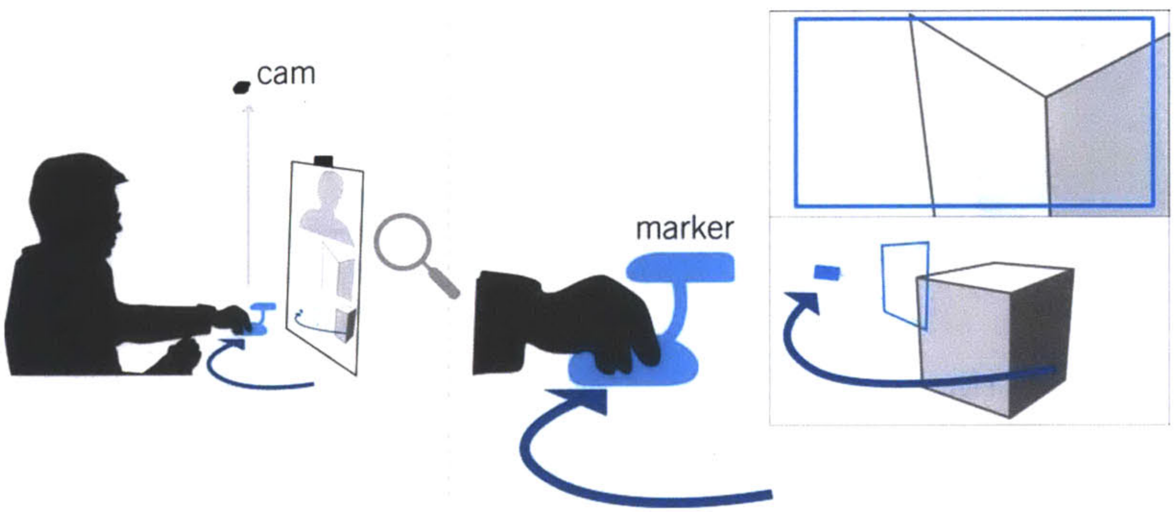

Figure 35: Illustration of tangible camera controller usage and viewport change

To assist in quick review of the archived annotations in multiple locations, we have also integrated animated transition mode that uses button inputs into the AnnoScape's environment. In this preview mode, participants can spatially navigate through existing virtual workspaces in smooth animation on a predefined trajectory created by the collaborator or themselves. The preview mode is yet to be evaluated, and we plan to explore this further in future work (Fig. 36 and 37).

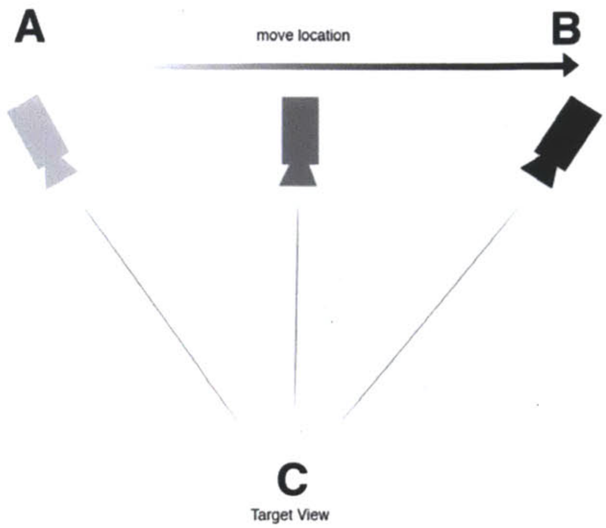

Figure 36: Navigation mode through smooth animation between captured viewports

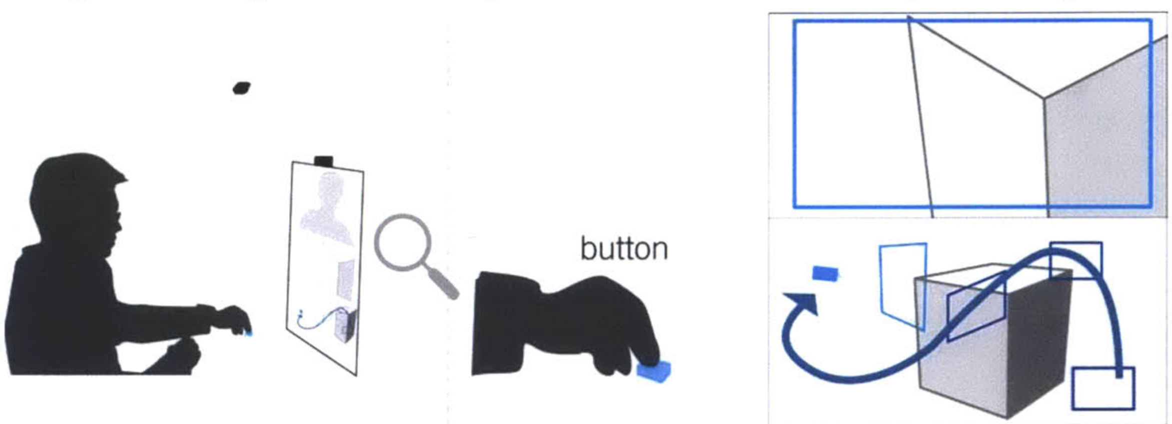

Figure 37: Illustration of preview mode using smooth animation

3.2.3 Annotation

AnnoScape's annotation feature is driven by the spatial video overlay technique applied in the 3D virtual workspace. A key concept of spatial video overlay is allowing users to superimpose the live video image from the desktop (e.g. their hand gestures, freehand drawing, paper print, physical material or models) onto any arbitrary locations in the virtual 3D workspace from a specific point of view. Both freehand and physical annotations can be spatially arranged onto the 3D scene. The content from the physical workspace is extracted from the foreground and is overlaid on the captured focused viewport as a legible live image.

The live video overlay can be viewed from both focused viewport and the reference view. For example, the remote collaborator can see the other participant's live video annotation session through the focused viewport while navigating in the scene and while seeing the thumbnail of the live image in the reference view on the specific location where the annotation is taking place. The live annotation session can be shared when the focused viewports are tightly coupled by snapping the individual virtual camera proxy to the other collaborator's location.

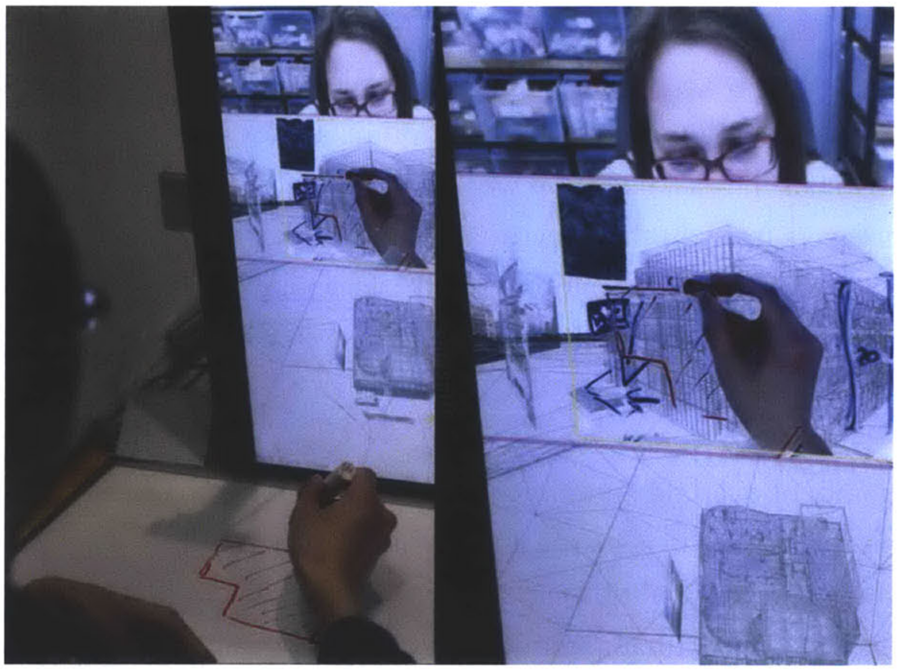

Figure 38: Annotation feature

In the annotation mode, the participant can use traditional drawing tools, such as pen and white paper. In the current setup, we are using a marker and a horizontal white board. By default, the system extracts non-white textures and materials such as the ink, hand, and physical object from the white foreground to provide a clean and legible 2D image. Through simple image processing, every white part from the desktop scene becomes transparent (Fig. 38). For example, if the participant wears a white glove, the hand becomes invisible inside of the virtual scene (Fig. 39). The system configuration also allows participants to spatially overlay live 2D images of physical materials as well as physical models onto the Focused Viewport.

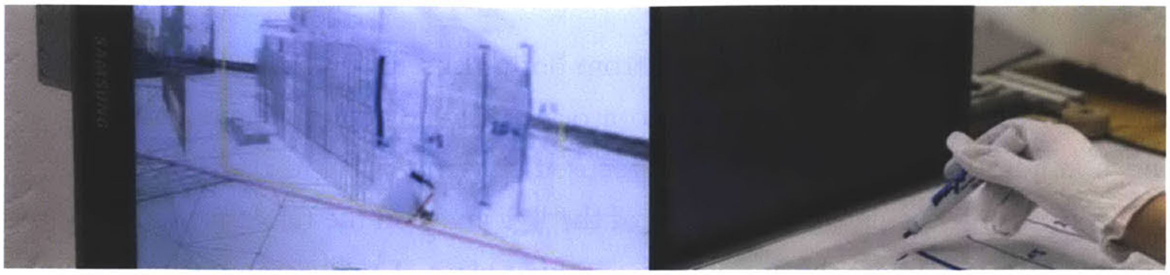

Figure 39: Making the hand transparent using white glove

3.2.4 Image Legibility Control

AnnoScape is designed to help remote participants to collaboratively review 3D digital information by generating shared annotation sessions in the virtual 3D space. Our system allows users to create, capture, and extract drawings or scenes on the individuals' physical desktop workspace in high quality and to display the 2D content layers freely in the 3D virtual shared workspace. As a 3D reviewing platform with a video overlay annotation interface, our work encompasses the fundamental process of shared-drawing media. Given this, we had to consider the following issues:

1) The legibility of overlaid annotation images decreases when too much information is clustered in a shared annotation workspace.
2) Also, when there are multiple users, identifying the owner of the captured images on the live overlaid annotation area can be difficult.

We considered managing or archiving the shared contents as essential interactive features in our system. To provide legibility of the live desktop video overlay on the shared digital information, we explored image-processing techniques to edit the visual appearance of the content, such as the color and transparency, based on the categories of the materials, drawing activity, ownership, and user's intent. We applied methods such as grouping, archiving, adding, and displaying the information in legible form in the virtual shared workspace. The proposed interaction technique is useful for managing the collaborative contents as well as for achieving a high degree of clarity and legibility of the layered information (Fig. 40). In the AnnoScape system, the associated contents that can be overlaid onto the

3D information exist as layers with live or recorded images of printed materials, photos, annotations, or drawings.

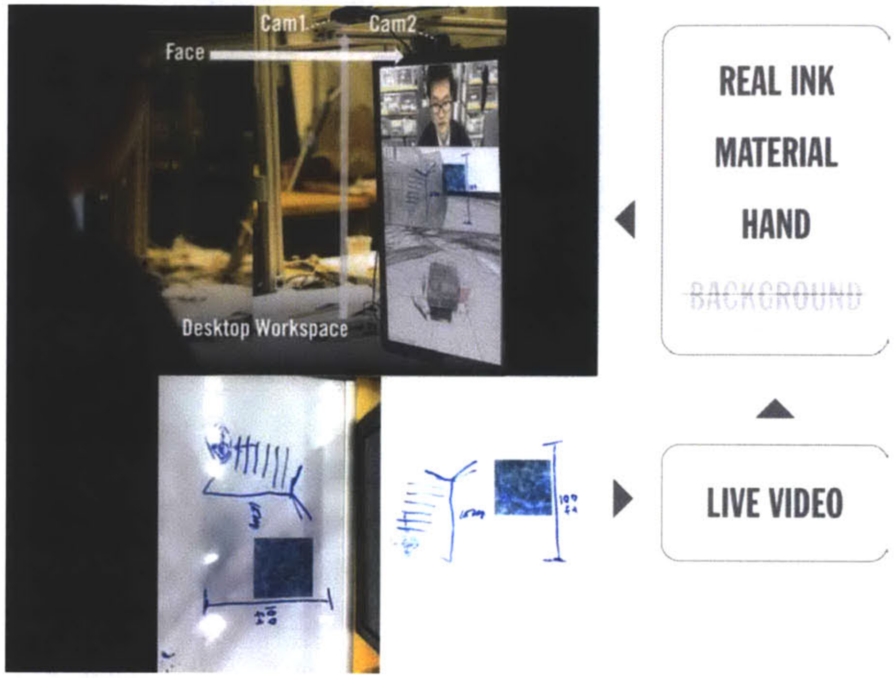

Figure 40: An example of image legibility control by extracting the background.

3.2.5 Work Modes

There are largely three types of work modes in the AnnoScape system based on how the viewports are tied to one another: 1) Individual Review mode is when a user works asynchronously on the 3D project; 2) Loosely Coupled Real-time Collaboration is when the remote collaborators are logged into the system, synchronously working on separate parts of the shared 3D data; and 3) Tightly Coupled Real-time Collaboration mode is when each viewport of the participants is synced. The following three diagrams (Fig. 41, 42, and 43) introduced in the sub-section scenarios illustrate the infrastructure behind the AnnoScape system that supports the individual review mode (1) and cooperative review modes (2, 3).

(1) Individual Review Mode Scenario

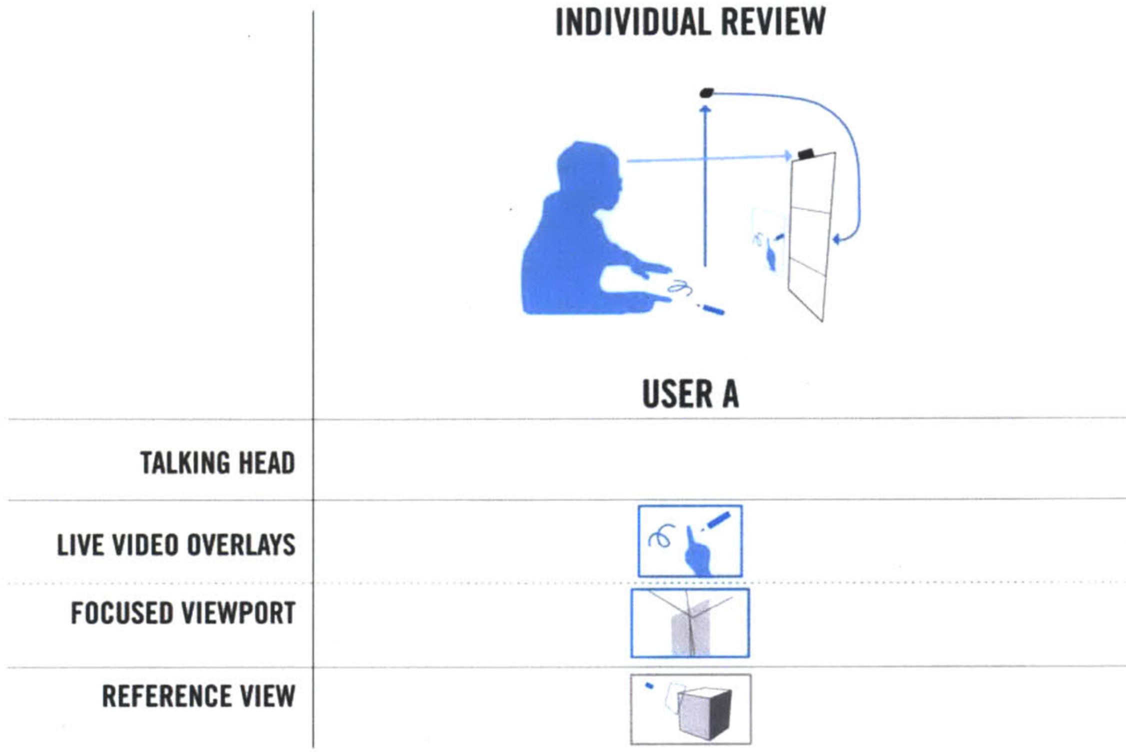

Figure 41: Individual Review mode

Scenario: Designer1 logs on to the AnnoScape system and starts to review the imported 3D model of the E14 building. His collaborator created the model using conventional CAD tools and sent him the file over night. Designer1 navigates through the 3D data by using his Tangible Camera Controller. By default, the interface is divided into three sections: IPS (inter-personal space), focused viewport, and the reference view (Fig. 5 and 41). The IPS is not showing the live image of the collaborator's face, and Designer1 realizes again that the time zone of his remote collaborator, Designer2, is different from his. For important parts of the data, Designer1 captures the viewport and leaves freehand annotation inside the virtual 3D scene by using traditional drawing tools on his physical desktop workspace. Every time he leaves the annotation session, the system automatically archives the captured scene along with the annotation overlay. Designer1 checks that all of his annotations are spatially arranged as small thumbnail images on the Reference View. Designer1 continues to navigate to different scenes using his Tangible

Camera Controller. The Virtual Camera Proxy showing in the Reference View helps the navigation by providing him with visual feedback.

(2) Loosely Coupled Real-time Collaboration Mode Scenario
The remote participant can also login to the system and individually generate annotation sessions in multiple locations in the 3D environment. This group work mode can be performed asynchronously or in real-time. Loosely Coupled Real-time Collaboration is when the collaborators are both logged on.

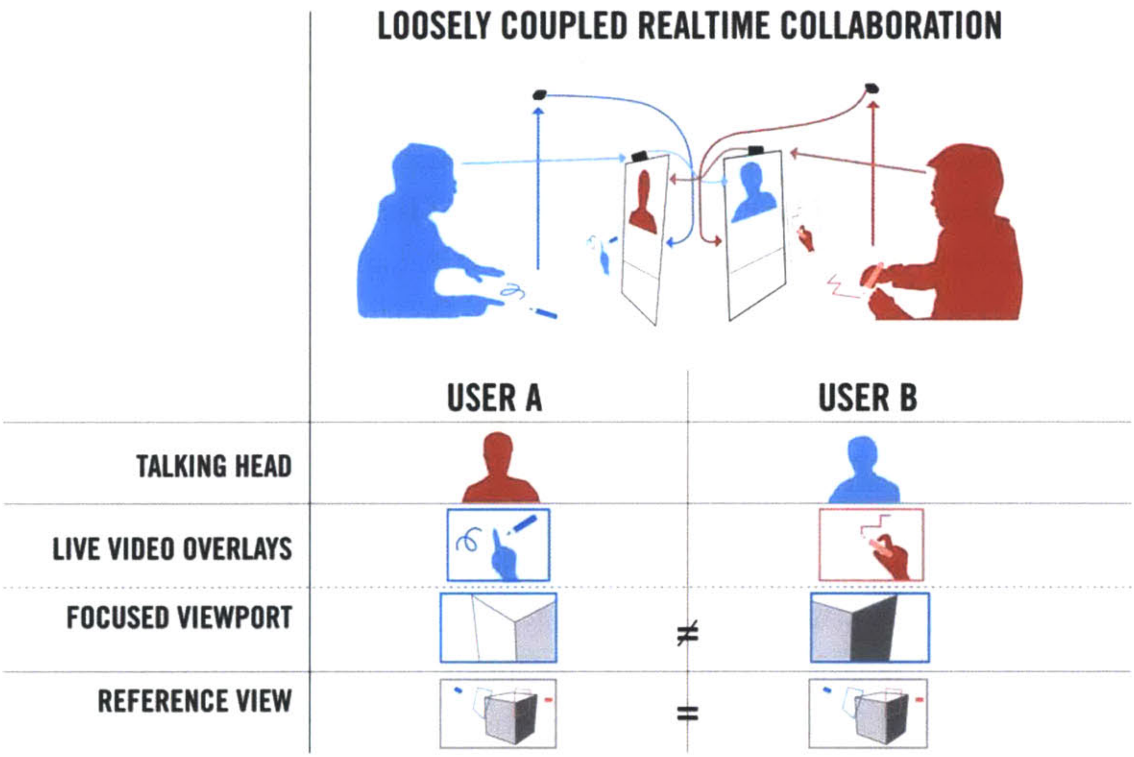

Figure 42: Loosely Coupled Real-time Collaboration

Scenario: After several hours, the remote participant, Designer2, wakes up and logs on to the AnnoScape system. Designer1 hears Designer2's voice and sees his collaborator in the IPS section (Fig. 6 and 42). The IPS supports the face-to-face communication by displaying the "Talking Head." Designer1 tells Designer2 to review the comments he left in various locations of the 3D data. Designer2 has two options to go through the archived annotations: The first is the default navigation mode that is based on the Tangible Camera Controller. The second is the quick preview mode, which allows the participant to spatially navigate through the archived annotation workspace in smooth animation. Designer2 decides to use the quick preview mode and clicks to each of the annotations. Designer2's virtual camera proxy smoothly flies between the workspaces in a predefined trajectory created by the collaborator, Designer1. Designer2 adds new layers of annotation on top of existing annotation workspace generated by Designer1. Designer2 switches to default navigation mode and continues reviewing Designer1's annotations by moving his personal Tangible Camera Controller. The scene on the Focused Viewport shifts, and the virtual camera proxy in the Reference View moves according to the Tangible Camera Controller's location. When the virtual camera proxy is near the archived annotation area, Designer2's viewport snaps to the specific location. Once Designer2 finishes reviewing all of the annotation, Designer2 works on the different parts of the 3D data. Meanwhile, Designer1, who is still logged on to the system, sees where Designer2 is leaving annotations through the Reference View. Designer1 can also revisit any of the captured annotation sessions generated by his collaborator, Designer2 and leave additional layers of annotations at any time.

(3) Tightly Coupled Real-time Collaboration Mode Scenario

When the remote participant joins the local participant's annotation session synchronously in the same location of the Focused Viewport, the shared viewport generates a strict WYSIWIS (What You See Is What I See) atmosphere that is suitable for a real-time shared review on the same 3D scene (fig. 7 and 43). This is the Tightly Coupled Real-time Collaboration mode. In this work mode, the interface can be switched between the default interface and the Tightly Coupled Collaboration Interface. The annotation area in the virtual 3D scene becomes a shared drawing surface that overlays both participants' physical desktop workspaces.

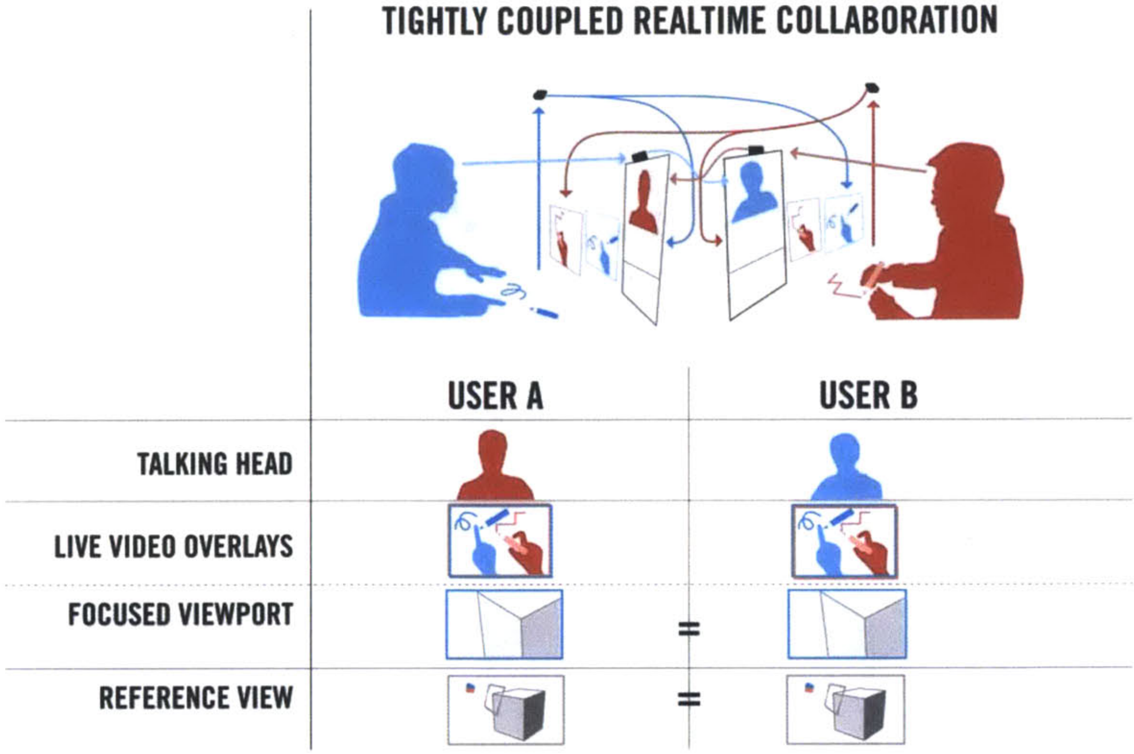

Figure 43: Tightly Coupled Real-time Collaboration mode

Scenario: From the Reference View, Designer1 notices that Designer2 is leaving an annotation on the entrance area of the 3D building model and decides to join the session. Designer1 moves the virtual camera proxy using his Tangible Camera Controller. As his virtual camera proxy moves closer to Designer2's annotation workspace, the live-image thumbnail of the captured desktop workspace is highlighted in yellow color. Designer1 captures the Focused Viewport and overlays his live desktop workspace onto the identical location in the 3D scene. Now, both of the remote collaborators' Focused Viewports are synchronized in real time. Designer1 shows physical material samples that can be used for the building's entrance area. Designer2 annotates on top of the captured material image while talking to his collaborator. Both hands of each user are showing in the Focused Viewport, and when the information becomes confusing, Designer1 opens the option panel to control the legibility of the created contents. Designer1 lowers the opacity of Designer2's content and continues the annotation from his individual scene.

3.2.6 Switching the 3D Data

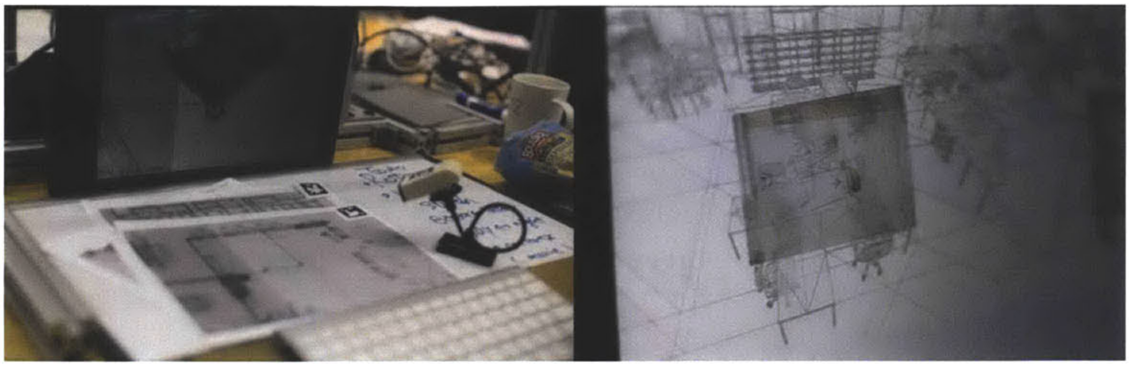

Figure 44: Switching to interior 3D model using print of the plan with AR marker

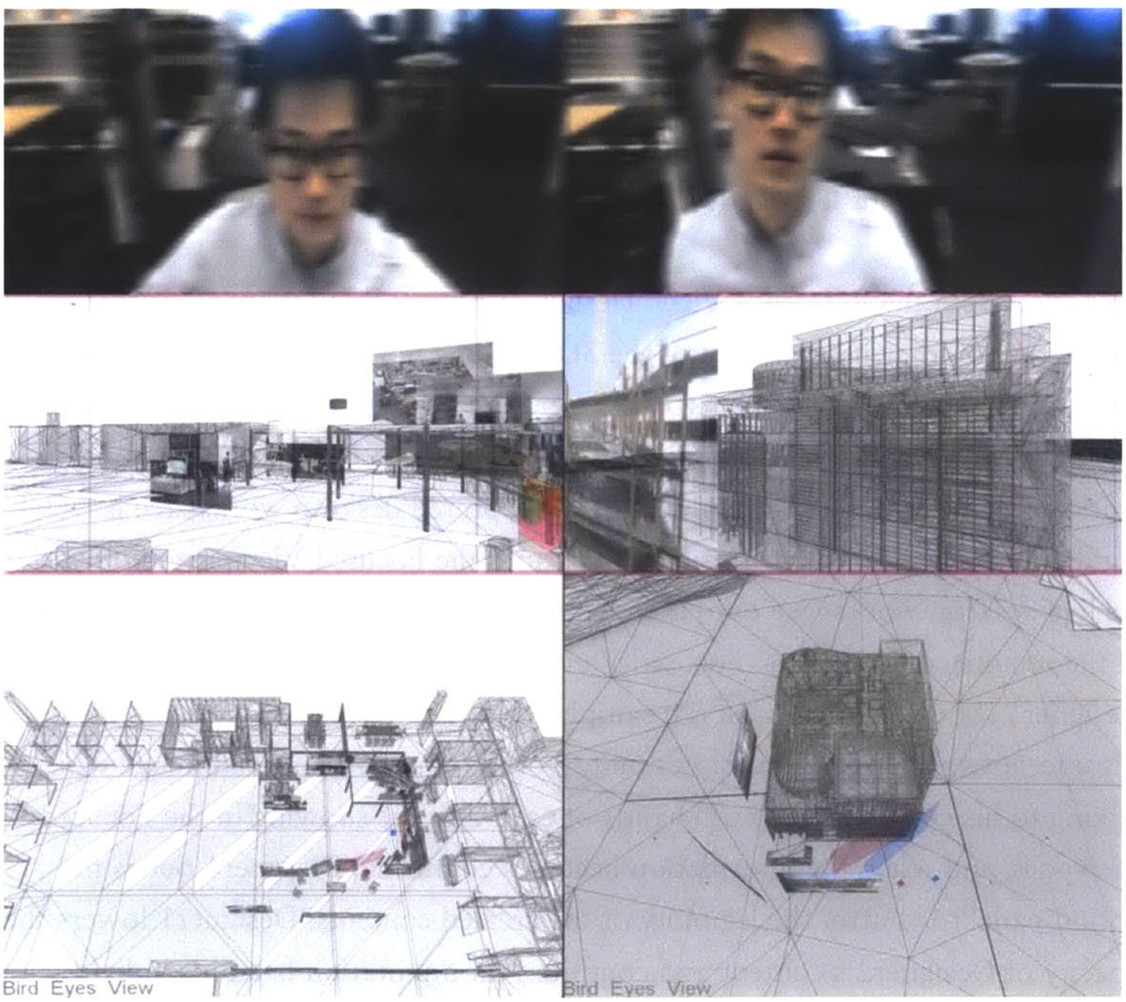

Figure 45: Switching 3D models. BT lab interior (left) and MIT E14 building (right)

Our system allows users to switch between the predefined 3D models by using a paper plot of the model with an AR marker (Fig. 44 and 45). The marker becomes the trigger to change the 3D model. When the 3D model is changed, the annotation sessions reset according to the 3D scene. In the current AnnoScape prototype, we import the 3D obj file of the Media Lab's E14-building model and the interior model of the BT lab. Users can switch to different scenes using the paper print. We embedded photos and videos from the actual physical location of the building area and spatially arranged the associated images in the virtual 3D data. For example, we included the time-lapse video of the building's construction site in our demo set-up and spatially arrange the video according to the actual location where the video was shot. Users can navigate to the scene and see the 3D model building being aligned with the video's architectural structure.

4 IMPLEMENTATION

The goal of this thesis is to provide the system design for collaborative remote 3D review using the video overlay technique. Rather than focusing on novel technical inventions such as a faster image-processing network, we focus our contributions on exploring new interaction techniques and evaluating our design through preliminary user study. We implemented the prototype system based on our design concept and guidelines. The current prototype simulates a situation in which each collaborator joins the annotation session from a different location by running an application written in C++ using the openFrameworks library in Mac Mini. We use the prototype as a platform to explore our design interactions.

4.1 HARDWARE

For each station, our system requires one monitor and two web cams connected to a computer installed on a traditional desktop workspace. The AnnoScape prototype system has two inputs from each user's site: a video source for user's face (Video 1) and an overhead mount video source for the user's physical desktop (Video 2). In our current setup, the pair of cameras is integrated in two stations (Fig. 47). In order to achieve the viewpoint manipulation, we use the Tangible Camera Controller (Fig. 46). The Video 2 detects the locations of these camera props by the markers attached on each Tangible Camera Controller using ARToolKit library. The four cameras used in this system are commercial web cameras (up to 1080p resolution is available.) As the output, each site has a vertically situated display that shows the other user's head (IPS), her/his own virtual camera view (Focused Viewport), and the bird's-eye view (Reference View) to indicate the virtual location of the users.

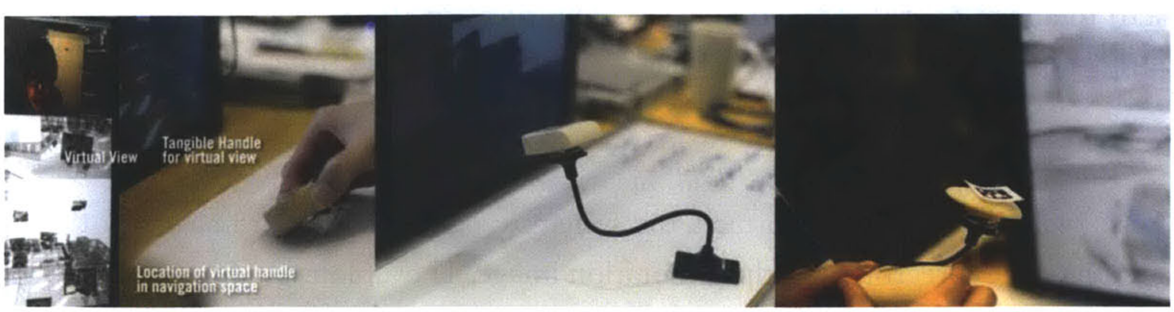

Figure 46: Design variations of the Tangible Camera Controller

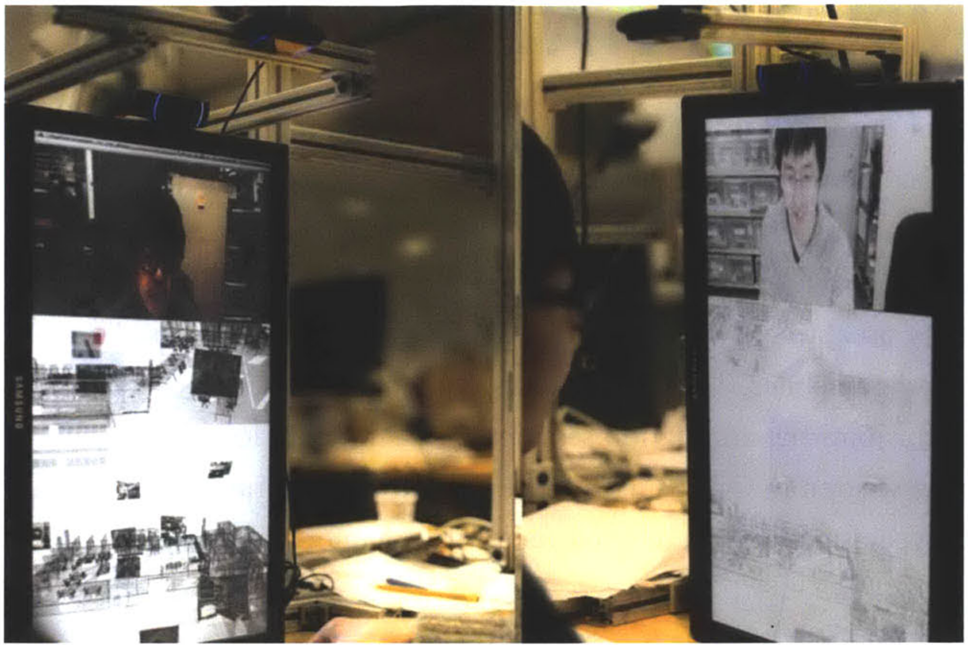

Figure 47: System hardware setup

4.2 IMAGE EXTRACTION FROM PHYSICAL DESKTOP

A background subtraction technique is a widely used approach for various Human-computer Interaction (HCI) applications (Elgammal et al., 2000). In order to make a clear and crisp annotation on the virtual 3D space, we implemented an image extraction feature. This is done through the following steps:

(1) Capturing the White Area

First, the system captures the white desktop as the background image before the annotation is performed. This step is for the white balance calibration. We place a whiteboard or an empty white sheet paper on the physical desktop workspace. Background images are automatically captured when users make a snapshot and start to annotate at that location. The calibration can also be done manually by pushing the delete key.

(2) Capturing the Annotation Image

Once the annotation is performed, the system recognizes the color images that are on top of the foreground as the annotation image. The annotation image can be physical ink, a paper plot of textures, hands, and other traditional desktop tools.

(3) Creating a Subtracted Image

Then, the system makes a subtracted image by using the annotation image and the background image.

(4) Binarizing the Subtracted Image

The next step involves binarizing the subtracted image and changing the alpha value of the pixels, which is below a certain threshold. Any images that have the same brightness as the background image are considered to be white. The white area of the captured image pixel is set to be transparent. This process extracts the annotation image from the foreground.

4.3 3D MODEL

For our prototype, we generated two scenes of the 3D models of an existing building, which was the E14 Media Lab building (Fig. 49). In current setting, users can switch the 3D models by placing the printed floor plan with a visual marker.

1) Exterior: We set the 3D model of a building and drew it with wireframes and faces. Pre-defined photos and video were placed around the building to situate the scene with useful information related to the architectural structure.
2) Interior: We set 3D model of the third floor's interior structure and drew it with wireframes and faces. Pre-defined photos captured inside the actual building were spatially arranged particularly in the relevant locations within the interior space to help users have a better sense of the site.

Parts of the ARToolKit library are used for indicating the visual marker as an input source. Once the marker is recognized, the system loads the 3D model and generates a virtual space for annotations. Currently, we are importing obj files but this can be any 3D format that can be read by openFrameworks (Fig. 48). We used Rhino to clean up and to convert existing 3D data to obj files.

Figure 48: 3D variation concept sketch (architectural images provided by S. Chang)

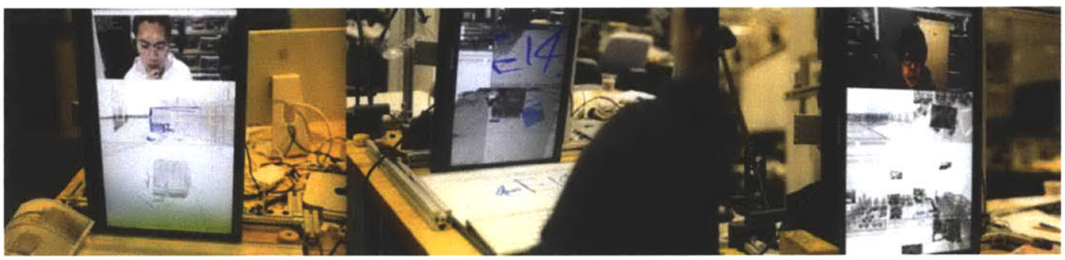

Figure 49: 3D model variations implemented

4.4 TRANSITIONS

For our prototype, we generated other features, such as preview mode, which allows users to smoothly shift to the various archived annotation workspaces. In the current system, we utilize keyboard input for some of the functions such as tilting the model or showing the interface for legibility control options (Fig. 50). Although the keyboard input can be replaced with custom tangible input controllers for this research, we focused on the system design for the fundamental platform for remote 3D review. The iterations of the design will be covered in the "Explorations and Future Work" chapter.

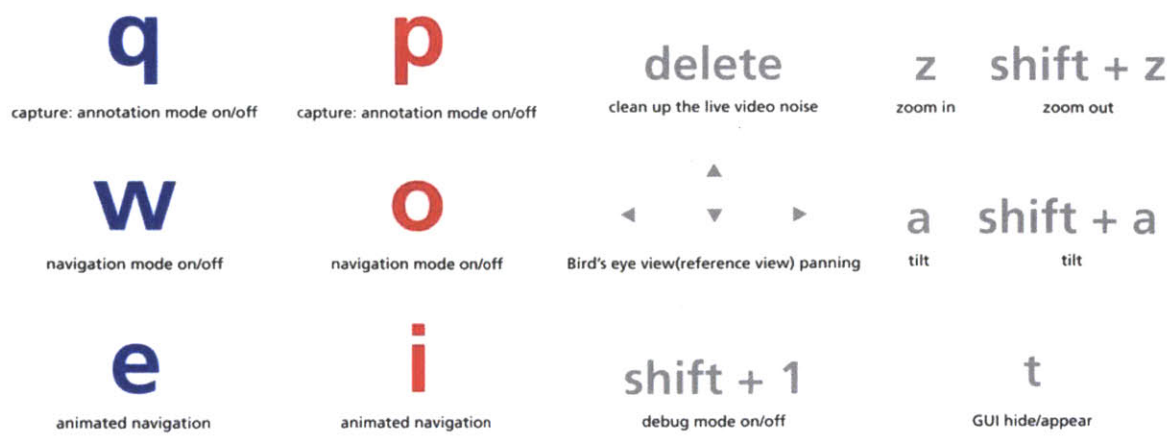

Figure 50: Input and features

We simulated two remotely connected machines and conducted our research in a local environment for this current prototype; however, the concept can be easily extended to a fully remote system using the two machines and a network.

5 EVALUATION

5.1 PRELIMINARY USER STUDY

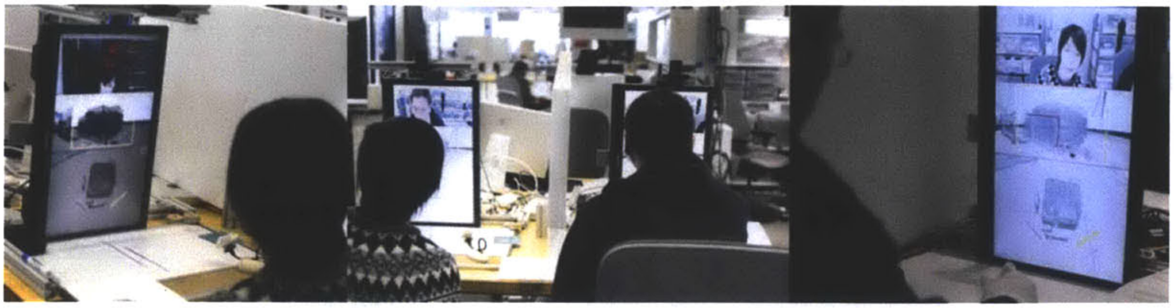

Figure 51: Setup of AnnoScape prototype for user study

We evaluated our system by conducting a preliminary user study with eight participants to investigate how the users can go through the 3D reviewing process using our system (Fig. 51). In addition, three architects were asked how the system could be used in their work. Our focus was especially on the ways in which users can annotate on the 3D data using conventional desktop work settings and investigate how a live video image of hand gestures captured in the virtual 3D space can benefit the collaborative review. Among these eight participants, two of them were architects, and one was an industrial designer who had experience with reviewing and annotating on the printed image of a 3D architectural model, while the rest had never experienced such a reviewing process.

We randomly chose two users as a pair for each study and gave them the tasks. First, we went through the feature walk-through with each participant. The features we introduced in the walk-through session were the face-to-face communication using IPS, navigating in the digital 3D space using tangibles, and locking into a position to capture the Focused Viewport. Lastly, we explained how the live image of the desktop workspace could be shared in the virtual 3D scene.

(1) First Task: Navigation and Capturing the Focused Viewport
Individual participants tried out all of the features in our system. We showed a hard copy of a scene (Fig. 52), and then, we asked the participants to navigate to the particular location and capture it. This was done in the Individual Review

Mode. All of the participants manipulated the Tangible Camera Controller and were able to successfully navigate to the scene.

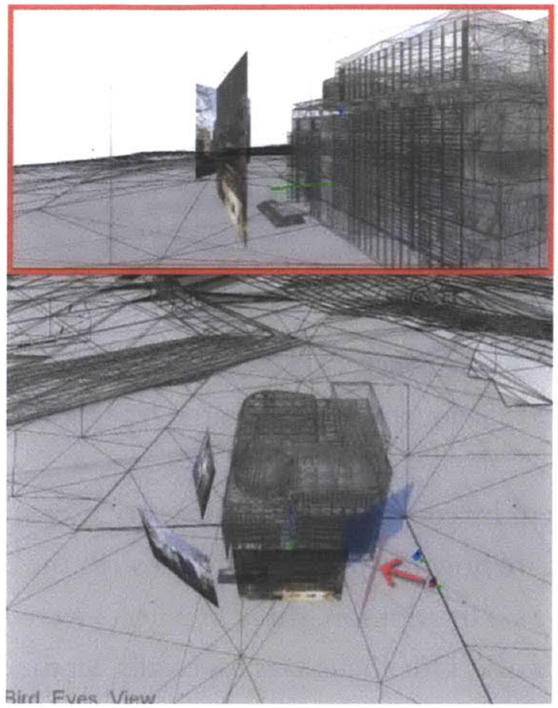

Figure 52: The participants were asked to navigate to the scene shown above using the Tangible Camera Controller. Once the participant found the scene, he/she were asked to lock into the position and to transit into the annotation mode.

The participants rated the fluidity of interaction and the ease or difficulty of completion for each task. From this experiment, we evaluated how difficult it was to approach and capture the particular scene.

(2) Second Task: Annotation in 3D Space

An individual participant annotates information on specific parts of the 3D model, and the other user puts that value into a column. We showed an example of simplified architectural annotation to the first participants and asked if they could annotate the height of the 3D model's component, then we asked the second participants to navigate to the scene and to read the annotation that the first participants left. After the tasks, the two pairs of participants had free-style discussions using verbal and annotation-based communication.

(3) Third Task: Annotation and Hand Gestures

The participants work at a synced Focused Viewport collaboratively using freehand annotations, a display of sample material prints, and finger pointing. For the third activity for the pairs, in the Tightly Coupled Real-time Collaboration mode, we had at least one participant initially cover her/his hand with a white glove and make the hand invisible within the Focused Viewport. (In our system, any white image becomes transparent. This is used for the background subtraction.) Next, we had the participant uncover her/his hands so that the collaborator could see the freehand gestures again. We evaluated when the hand gestures were useful during the review session.

(3) Fourth Task: Physical Annotation and Finger Pointing

The participants captured material images from the desktop workspace and used them as annotation content. Participants were asked to utilize the paper plot of the material sample to communicate the possible textures that could be used for the design. In our setup, we had the participants collaboratively decide the exterior façade and entrance stairs using annotation, physical texture capturing, and freehand gestures.

Figure 53: The tasks included navigation, capturing, annotation, finger pointing and overlaying live steaming image of physical materials.

In summary, in the exterior 3D model-viewing situation, the first participant was asked to begin reviewing the model and to find the entrance; then, the participant was to put the annotation that indicated the size of entrance (e.g. width = 5m). Also, we had a second participant start to review the identical model and the annotation sessions that had been created (Fig. 53). The second participant was also asked to annotate around the same Focused Viewport (e.g. captured image of the entrance area). We had the participants have casual conversations about the 3D information while having each participant annotated on top of the other user's annotations.

5.2 RESULT

To learn whether the video overlay was useful for annotations in 3D, we observed the usage of physical objects and materials on the desktop workspace (Fig. 54). A two tailed paired t-test with a significance value of $p = 0.05$ was used for questions rated 1-7. A similar test, sign test, was used for binary responses of preference questions. There were no significant differences for white gloves versus no white gloves. All questions related to material samples were significantly better when the material samples were used. The bar graph is user preference of materials samples versus no material samples (Fig. 55, 56 and 57). Everyone preferred using material samples.

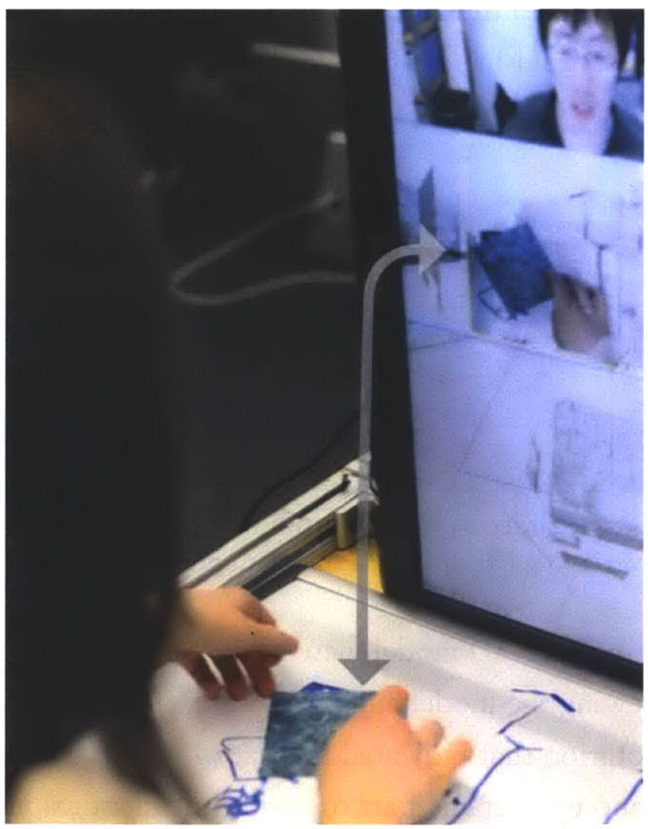

Figure 54: Material sample usage example

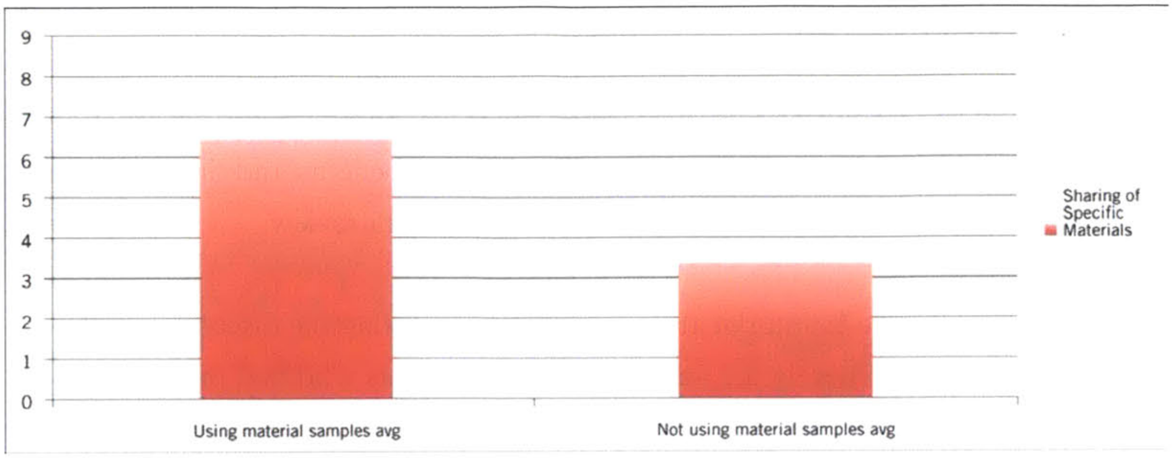

Figure 55: Overall usability comparison (p-value 0.0123).

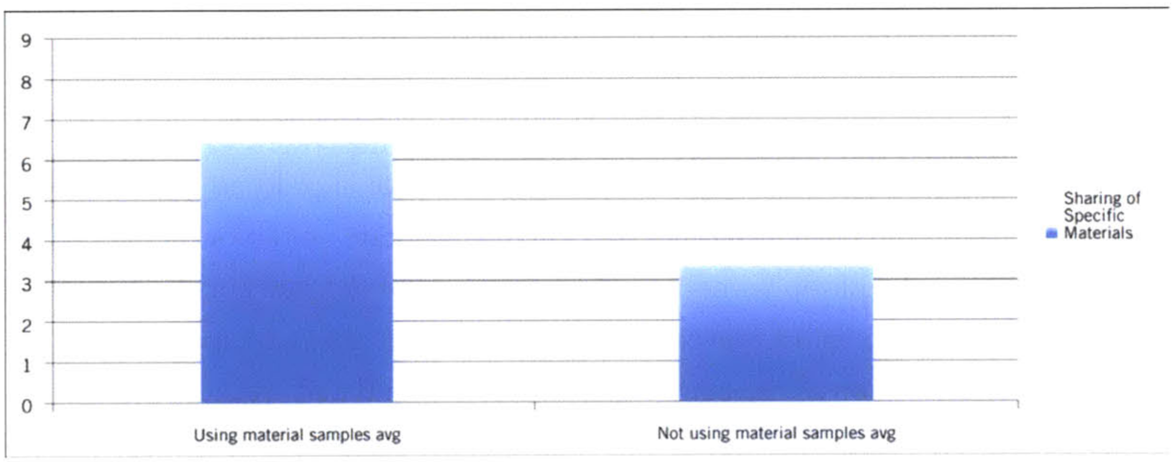

Figure 56: Sharing of specific materials (p-value 0.0046)

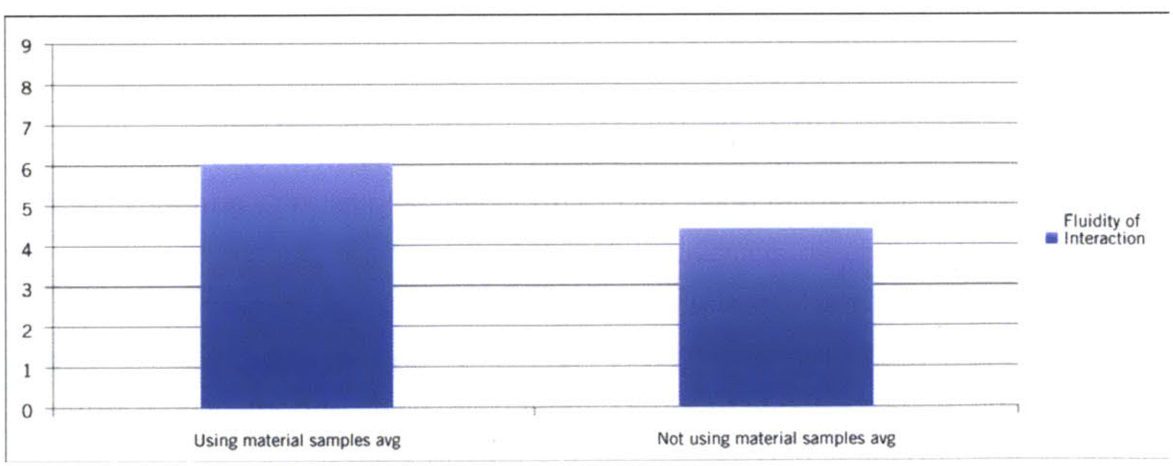

Figure 57: Fluidity of interaction (p-value 0.0062)

5.3 EVALUATION

All eight participants successfully performed the tasks. Based on our observation and the survey we conducted, we learned the potential benefits and limitations of video overlay-based annotation for collaborative 3D data review.

Hand gestures. People found the ability to see hands during the discussion to be helpful; however, for merely the annotations, participants who had professional design backgrounds preferred to see only the strokes without the hands covering the 3D scene. Annotating, pointing, and moving objects happened almost simultaneously during the discussions, and in general, participants all thought that having the hands visible felt more natural than having them invisible.

Indirect Drawing. Two participants felt that annotating on top of the 3D model precisely through indirect drawing would require practice in order to get used to it. This is because in our current setting, participants annotate on top of the horizontal physical desktop without a reference image of the viewport being displayed on the desktop. However, we learned that once the participant was able to locate the orientation of the marking of her/his pen in the 3D scene, he/she could annotate successfully with ease.

Possible Solutions. A professional architect suggested that the difficulty with indirect drawing could be resolved drastically by relocating the Focused Viewport. Having the Focused Viewport displayed on the bottom of the AnnoScape closer to the desktop area would allow users to see the video overlay and the desktop workspace simultaneously in one glance. This would help users to annotate with better precision. Two users mentioned that if there were less latency, the annotation would be much easier. In our study we did not demonstrate the features of the dynamic interface to focus on the key benefits of live video overlay. Regarding the reference image for direct annotation, while options such as using overhead projection may have been a possible solution, it also had the potential of generating new problems such as casting harsh shadows on the workspace during the annotation session and requiring that existing work settings be drastically changed. Using a tablet with a reference image as an input source is being considered as an alternate solution, as it would allow additional collocated users to participate in the collaborative remote review sessions. However, with a single tablet-based application, sharing physical annotation or hand gestures would have been difficult to achieve.

Background Subtraction. Another comment was that the background subtraction was not being performed perfectly when unwanted objects were in the foreground. Because our system automatically calibrates the background by default, unwanted objects on the desktop would sometimes create markings when the background was subtracted. We implemented a one-click manual calibration feature that would enable users to achieve clean background subtraction.

Professional Review. The three architects gave positive feedback from pragmatic points of view during the interview, and they mentioned possibilities of extending the system platform for surveillance application or underground pipeline design review. The architects found the system to have great potential for internal design review sessions. However, according to the architects, the system was not considered suitable for the final presentation to the client, mainly due to current 3D rendering settings. For commercial usage, replacing the 3D obj file with fewer polygons or importing the 3D data with simple mapping could improve the visual quality drastically. Although the scope of this particular research focused on evaluating the interaction techniques using fast 3D renderings, for future work we plan to improve the visual quality of the 3D data.

Spatial Video Overlay. A majority of the participants found the spatially arranged 2D information, such as photos and videos in the 3D scene, to be engaging, and they showed interest especially in the possibilities of expanding the video overlaying technique in the 3D space.

Possible Applications. Currently, other than the desktop video overlay, we are using existing pre-processed videos. However, for future work we plan to integrate a real-time live video feed from various physical locations into the 3D digital model. For example, the on-site construction process can be shown inside the virtual 3D model in the precise location.

5.4 DISCUSSION

5.4.1 Indirect Drawing Experience

The most frequently mentioned limitation of the AnnoScape platform was the indirectness of the annotation input. Through our observation, we realized that participants with a professional design background tend to trace the 3D image in the scene during the annotation session. For these participants, especially in the initial stage of the annotation, the indirectness of the drawing was problematic. The participants suggested methods that could potentially upgrade the drawing experience, such as improving latency and reconfiguring the interface. We have implemented the dynamic interface reconfiguration feature customized for the work modes of the AnnoScape platform. However, the relation between the interface configuration (Fig.58) and the drawing performance is yet to be evaluated. Although a couple of participants mentioned minimizing the latency as a solution, most people did not feel that latency was as an issue. Because of the scope of this research, rather than concentrating on making the program run faster with ultra low latency, we analyzed our preliminary user study and observations to focus on the new interaction techniques that we can apply to improve our system design.

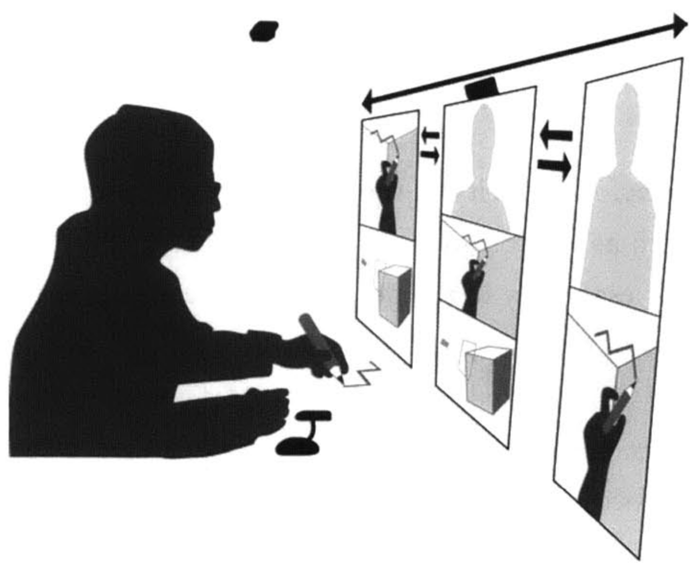

Figure 58: AnnoScape dynamic interface configuration options

5.4.2 Application of Traditional Drawing Assisting Tools

During our observation, we noticed that one of the users with an architectural background first marked dots according to the 3D scene and used them as reference point for the annotation. Later, we realized that all of the users were able to perform the freehand annotation with ease once they established a reference area during the annotation session. We decided to provide traditional drawing tools used for technical sketching, such as rulers, set square, and protractor onto the physical desktop workspace. Through informal user testing, we learned that the participant could use drawing assisting tools, such as a ruler, to create a physical reference point for the annotation. Although having drawing assisting tools resolved the inaccuracy of the indirect drawing drastically and increased the speed of the performance, having too many tools on the desktop workspace made it complicated to share important information with the collaborator. For example, the live image of the annotations with physical ink, material samples, and technical sketching tools seemed to create a cluttered scene with less legibility. This could be resolved by partially covering the drawing assisting tools with white color to make them appear transparent in the live video overlay image. This method was explored during our preliminary user study when the participants were asked to cover their hands with white glove.

6 EXPLORATIONS & FUTURE WORK

6.1 EXPLORATIONS

In this section, we share a set of research explorations as well as the design iterations related to our final platform. Our design process includes studies on gaze for viewing 3D objects (Fig. 59), experiments related to sketching that can be extruded in 3D, and the initial prototype design that merges IPS, the desktop video overlay image and the 3D data using network connection. Although many of the design experiments were not directly applied in the final AnnoScape platform, all the prototypes explored different aspects of remote 3D data review and provided technical resources and design guidelines for the final implementation of our work.

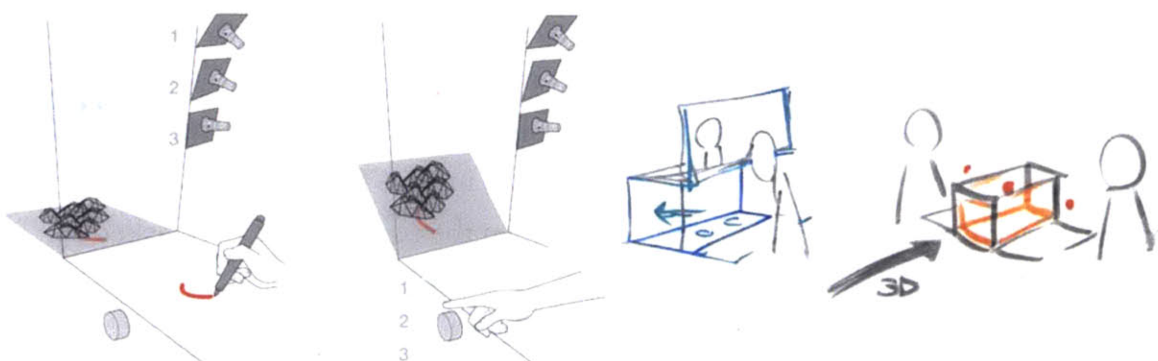

Figure 59: Idea sketch for system design to support the types of gaze for reviewing 3D

6.1.1 Gaze Study for 3D Contents

We explored the characteristic of gaze for 3D contents in the real world and attempted to apply our findings in virtual 3D content review. This research has involved analysis of the behaviors that occur when people view 3D objects in the real world. Also, the research included work in software and exploring the design of the hardware (Fig. 60, and 61). Our initial attempt was to support the gaze awareness feature for 3D contents. Based on the analysis, we created a prototype to explore the dynamic gaze shifting. After creating the prototype, we realized that because of the virtual distance between the "Talking Head" and the 3D data and the issues of WYSIWIS, it was difficult to support the gaze awareness feature in a seamless way for remote participants. However, through this process, we were able to explore the interface design as well as the hardware configuration.

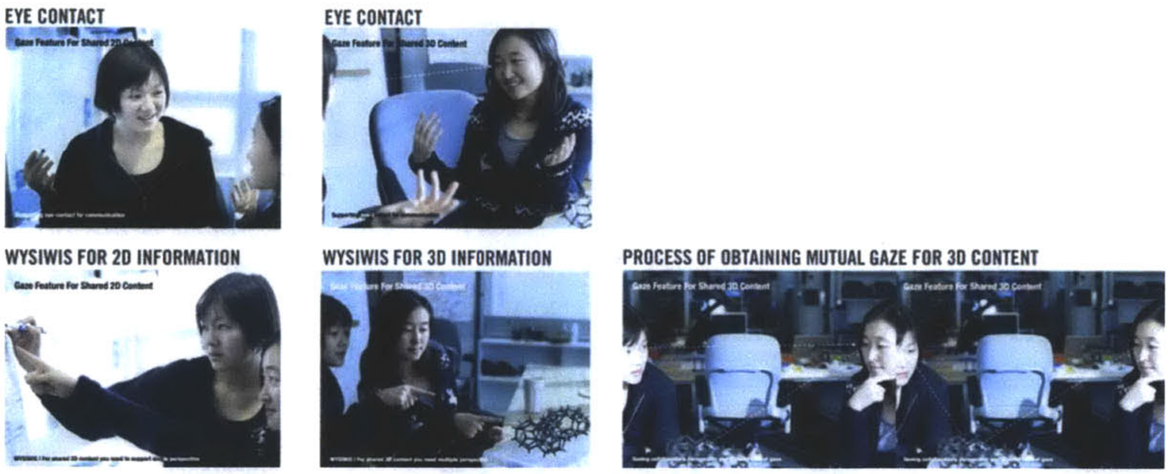

Figure 60: Types of gaze for reviewing 2D content (left) and 3D content in real space

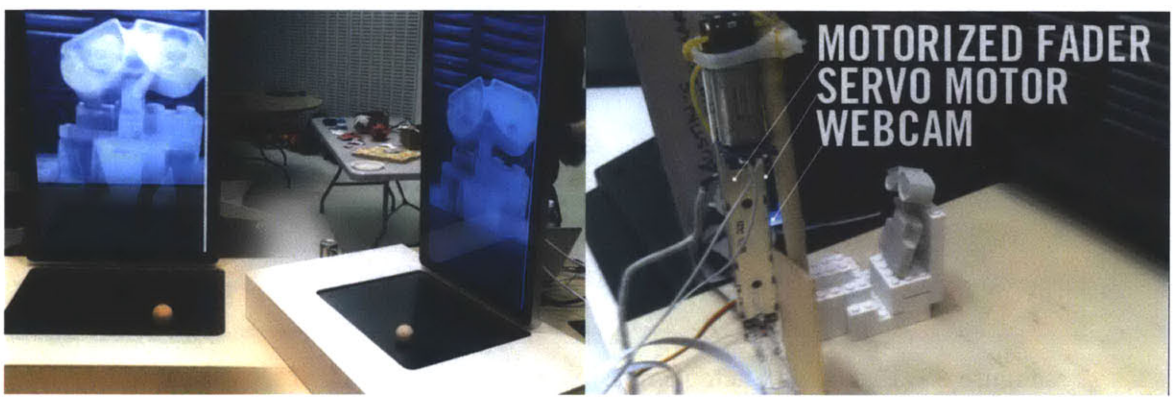

Figure 61: Hardware design iteration for dynamic 3D gaze experiment.

6.1.2 Ink with 3D Extrusion Feature

Three-dimensional sketching can be explored as a medium of expression. Initially, our system design included annotation in 3D that can be spatially arranged in z axis and annotation that can be extruded in to 3D (Fig. 62 and 63) Although this earlier prototype was useful for creating artistic images, in a pragmatic sense it was less useful compared to annotation in 3D. Moreover, creating a 3D form only with sketching was less effective than using state-of-the-art CAD software. Throughout this exploration we were able to define the scope of our research in regards to 3D annotation.

Figure 62: Extrusion Sketch working prototype.

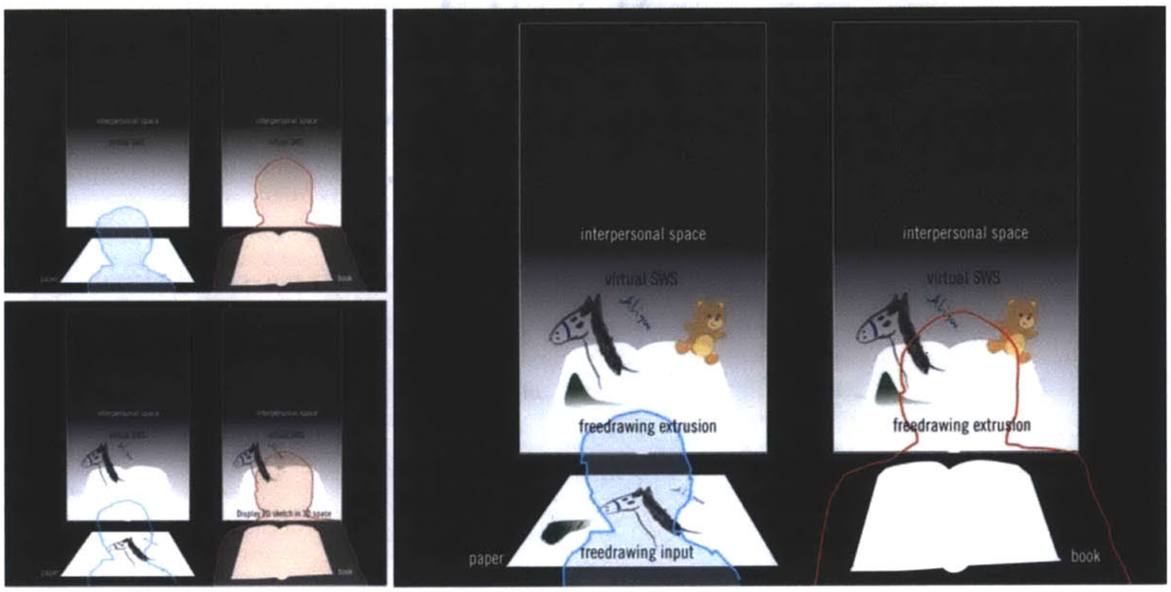

Figure 63: Extrusion Sketch's possible application scenario idea sketch.

6.1.3 Networked Shared Drawing

This project focused on creating the initial interface design for the AnnoScape platform. In order to build a shared platform for 3D content, we had to initially create a platform for shared drawing. During this process we were able to connect two separate machines via a network and successfully conduct our experiments. The platform was used for other iterative experimental designs (Fig. 64, 65, and 46).

Figure 64: Shared surface network connection and importing 3D data (top right).

In addition to these explorations, we implemented a gestural input feature—proximity sensing tangible objects—as a virtual proxy controller, which turned out to be not applicable to our current prototype setup. However, as a body of work, the design explorations provide useful guidelines for possible future applications.

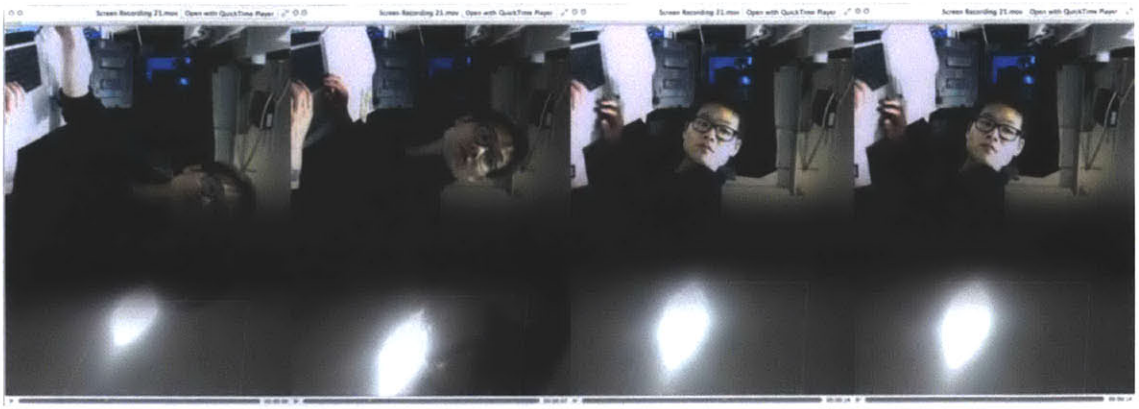

Figure 65: Interface design exploration to seamlessly connect the IPS with viewport area.

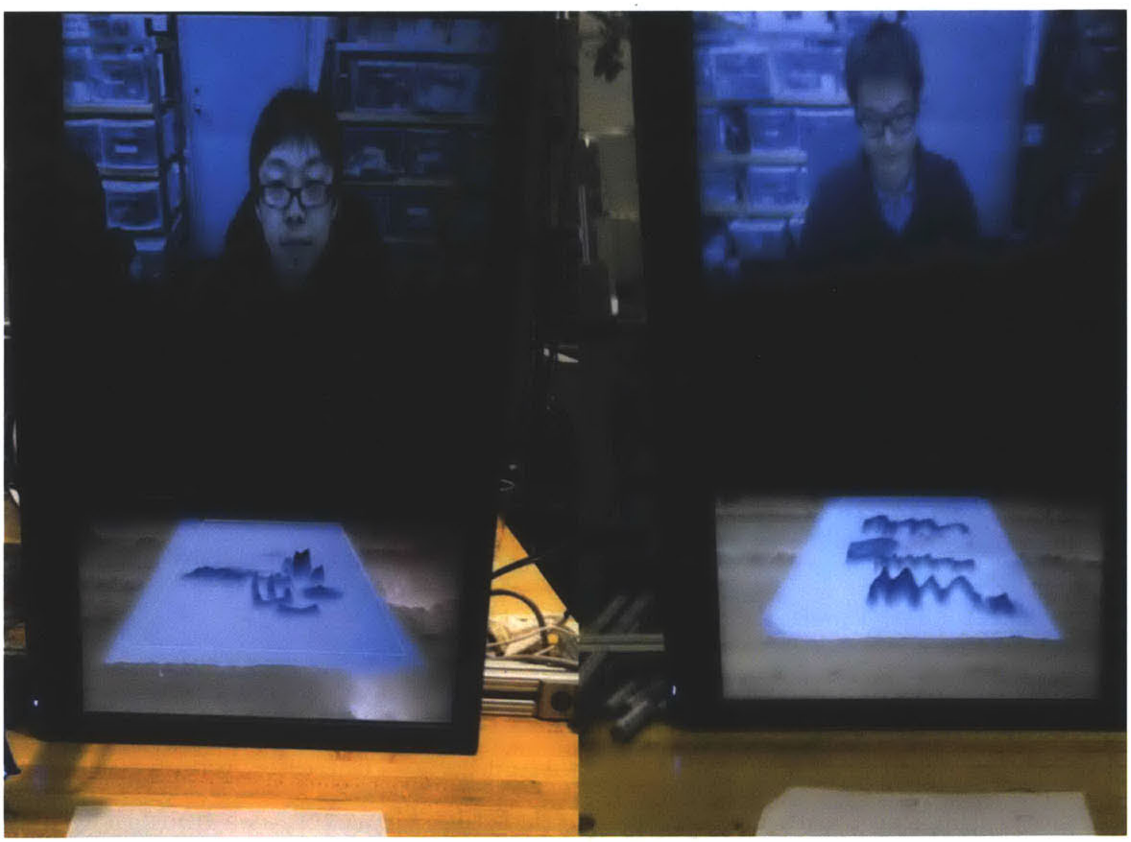

Figure 66: Interface exploration: prototype design with shadow masking.

6.2 FUTURE WORK

The AnnoScape platform was proven to have great potential for collaborative remote 3D review. During several events at the MIT Media Lab, people from industry gave extremely positive responses to our system design. Among many possible future directions, some of the most promising applications could be integration of a tablet-based interface for the AnnoScape platform and explorations with tangible media for assisting the annotations. The following concept images are related to the applications (Fig. 68).

In the tablet application, a user can have access to the 3D data by using either the desktop-based AnnoScape application or the tablet-based version (Fig. 67). This allows collocated users to scale up the participants, and those who do not have the desktop interface could also participate in the review session.

Figure 67: AnnoScape tablet application: reviewing 3D data concept image

Another possibility is combining the Second Surface project with the AnnoScape system (Fig. 69). As one of the contributors to the Second Surface platform, I realized the potential of using the real physical space for expressive spatial annotations. We could imagine a designer reviewing 3D data of an existing building using the AnnoScape platform leaving annotations in a specific part of the architectural structure. The collaborator with the tablet that is connected to the system then would walk to the physical location correlated to the virtual location and would see the live video overlay annotation in real space on the tablet screen.

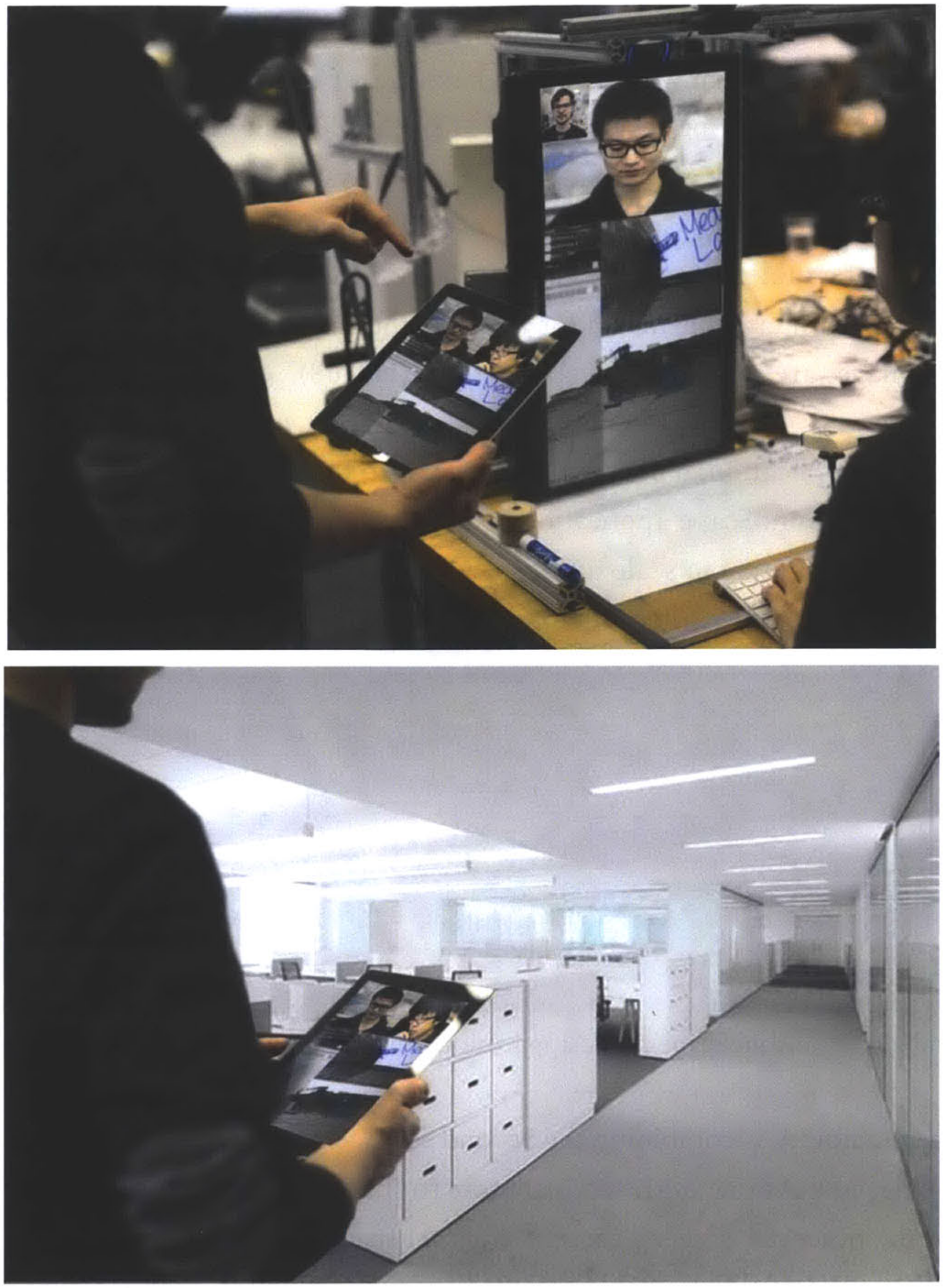

Figure 68: Tablet application for remote participant concept image

Figure 69: Combination of Second Surface and AnnoScape application concept sketch

The current AnnoScape system is a platform with fundamental features for collaboratively reviewing 3D information related to architectural projects. However, with the right design insights, we believe we can expand the project to potentially more diverse applications, such as geo-science or medical 3D review (Fig. 70). Also, we believe there is great potential in adding temporal dimension to the system. For example, recording the annotation session and archiving the live moving images may lead to more useful and interesting communications during the remote 3D collaborations.

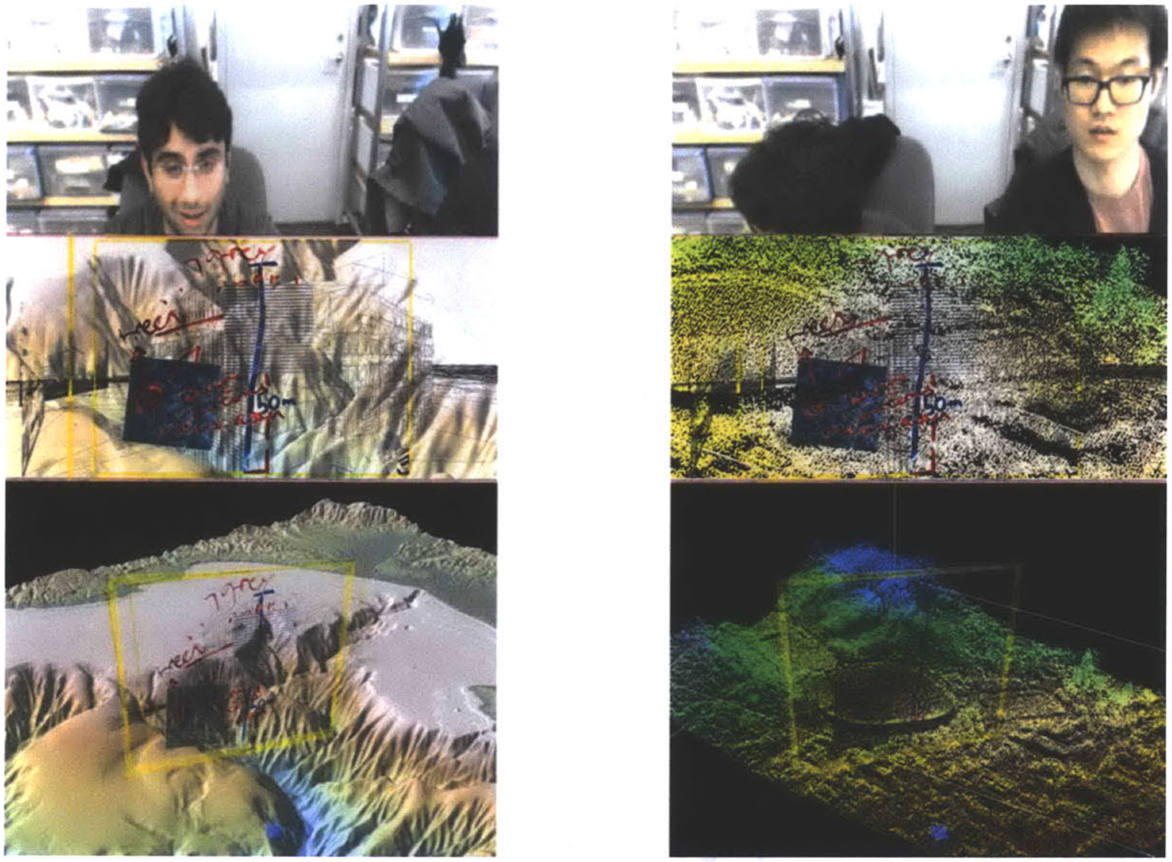

Figure 70: Geo-science data import example concept image

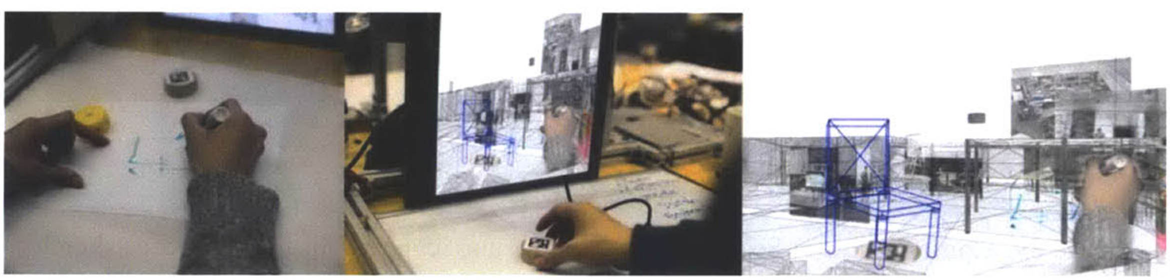

Figure 71: Tangible handle 3D model control example concept image

7 CONCLUSION

We presented the novel approach of AnnoScape for 3D data sharing for the collaborative review. Our approach supports real-time and asynchronous remote collaboration using the live video overlay technique in a virtual 3D workspace. We explored the possibilities of this system through two application scenarios using MIT Media Lab building and the BT lab 3D models, and a user study demonstrated that the collaborators could smoothly go through the process of reviewing and annotating on 3D data.

Our AnnoScape system introduces an interaction technique that preserves the traditional work style on the desktop for annotating on the virtual 3D data. By overlaying a live image of the physical desktop workspace, users can quickly annotate or use physical objects to visually support the communication during the collaborative 3D review. Our system aims to provide seamless workflow between the work done at the physical desktop with the virtual 3D reviewing, to allow designers to easily communicate ideas about the 3D data over a distance with the collaborator. Through a preliminary user study we show the benefits of using the live video overlay image for annotation in 3D, such as the ability to display hands and physical materials during the annotation session. We also introduce a set of design explorations to provide a guideline for potential future research.

We will investigate the further possibilities of configuration and the tangible approach in the near future. For example, we are planning to extend the system to allow users to reconfigure the furniture arrangement in the 3D model using graspable tangible models (Fig. 71). Achieving a seamless transition across a variety of work modes and enriching the representation of shared information are our long-term goals. This paper focused on the 3D data, three work modes (individual work, loosely-coupled collaboration, and tightly-coupled collaboration), and dynamic interface configuration. However, we would like to expand the system to support more complex data (e.g. temporal data) as well as participations of mobile users with wireless tablets, for example.

REFERENCES

Ajaj, R., Vernier, F., Jacquemin, C., & Cedex, O. (n.d.). Navigation Modes for Combined Table / Screen 3D Scene Rendering, 141–148.

Aliakseyeu, D., Subramanian, S., Martens, J., & Rauterberg, M. (2002). Interaction Techniques for Navigation through and Manipulation of 2D and 3D Data.

ARCADE: A System for Augmenting Gesture-Based Computer Graphic Presentations | SIGGRAPH 2012. Retrieved May 6, 2013, from http://s2012.siggraph.org/attendees/sessions/arcade-system-augmenting-gesture-based-computer-graphic-presentations

ART+COM : Terravision. Retrieved May, 2013, from the World Wide Web: http://www.artcom.de/en/projects/project/detail/terravision/

Benko, H., Jota, R., Wilson, A. D., & Inesc-id, V. (2012). MirageTable : Freehand Interaction on a Projected Augmented Reality Tabletop, 199–208.

Billinghurst, M., & Kato, H. (2002). Collaborative augmented reality. Communications of the ACM, 45(7). doi:10.1145/514236.514265

Digital Project | Gehry Technologies. (n.d.). Retrieved May 6, 2013, from http://www.gehrytechnologies.com/digital-project

Elgammal, Ahmed, David Harwood, and Larry Davis. "Non-parametric model for background subtraction." Computer Vision—ECCV 2000. Springer Berlin Heidelberg, 2000. 751-767.

Fitzmaurice, G. W., & Buxton, W. (1995). Bricks: Laying the Foundations for Graspable User Interfaces, 442–449.

Fjeld, M., Bichsel, M., & Rauterberg, M. (n.d.). BUILD-IT : a brick-based tool for direct interaction, 4(1), 205–212.

Forlines, C., Esenther, A., Shen, C., Wigdor, D., & Ryall, K. (2006). Multi-user, multi-display interaction with a single-user, single-display geospatial application. Proceedings of the 19th annual ACM symposium on User interface software and technology - UIST '06, 273. doi:10.1145/1166253.1166296

Hartmann, B., Morris, M. R., Benko, H., & Wilson, A. D. (2010). Pictionaire : Supporting Collaborative Design Work by Integrating Physical and Digital Artifacts.

Ishii, H. (1990). TeamWorkStation: towards a seamless shared workspace. Proceedings of the 1990 ACM conference on ..., (October), 13–26. Retrieved from http://dl.acm.org/citation.cfm?id=99337

Ishii, H., & Kobayashi, M. (1992). ClearBoard: a seamless medium for shared drawing and conversation with eye contact. Proceedings of the SIGCHI conference on Human .... Retrieved from http://dl.acm.org/citation.cfm?id=142977

Ishii, H., & Miyake, N. (1991). Toward an open shared workspace: computer and video fusion approach of TeamWorkStation. Communications of the ACM, 34(12), 37–50. doi:10.1145/125319.125321

Junuzovic, S., Inkpen, K., Blank, T., & Gupta, A. (2012). IllumiShare: sharing any surface. Proceedings of the 2012 ACM ..., 1919–1928. Retrieved from http://dl.acm.org/citation.cfm?id=2208333

Kasahara, S., Heun, V., Lee, A., & Ishii, H. (2012). Second surface: multi-user spatial collaboration system based on augmented reality. SIGGRAPH Asia

2012 Emerging Tehcnologies Retrieved from http://dl.acm.org/citation.cfm?id=2407727
Koutsabasis, P., Vosinakis, S., Malisova, K., & Paparounas, N. (2012). On the value of Virtual Worlds for collaborative design. Design Studies, 33(4), 357–390. doi:10.1016/j.destud.2011.11.004
Mckim, R. (1973). REFERENCE LIBRARY. Brooks/Cole Publishing Co.
MoMA | Front Design. Sketch Furniture. 2005. (n.d.). Retrieved May 6, 2013, from http://www.moma.org/explore/multimedia/audios/37/856
Moum, A. (2010). Design team stories. Automation in Construction, 19(5), 554–569. doi:10.1016/j.autcon.2009.11.007
Pablo Picasso, 1949 | LIFE Behind the Picture: Picasso "Draws" With Light | LIFE.com. (n.d.). Retrieved May 6, 2013, from http://life.time.com/culture/picasso-draws-with-light-1949/#1
Photosynth - Capture your world in 3D. (n.d.). Retrieved May 6, 2013, from http://photosynth.net/
Poupyrev, I., Tomokazu, N., & Weghorst, S. (1998). Virtual Notepad : Handwriting in Immersive VR HIT Lab , University of Washington.
Revit | Building Design| Autodesk. (n.d.). Retrieved May 6, 2013, from http://www.autodesk.com/products/autodesk-revit-family/overview
Rekimoto, J. (1999). Augmented Surfaces : A Spatially Continuous Work Space for Hybrid Computing Environments.
Stefik, M., Bobrow, D. G., Lanning, S., Tatar, D., Palo, X., Alto, P., Foster, G., et al. (1986). Wysiwis revised: early experiences with multi-user interfaces, 276–290.
Tangible Media Group. (n.d.). Retrieved May 10, 2013, from http://tangible.media.mit.edu/project/tether/
Tsang, M., Fitzmaurice, G., Kurtenbach, G., Khan, A., & Buxton, B. (2002). Boom {Inserting}Chameleon : Simultaneous capture of 3D viewpoint , voice and gesture annotations on a spatially-aware display THE LEFT, 4(2), 111–120.
Tyler DeWitt: Hey science teachers -- make it fun | Video on TED.com. (n.d.). Retrieved May 6, 2013, from http://www.ted.com/talks/tyler_dewitt_hey_science_teachers_make_it_fun.html
Wellner P., Freeman S., The Double Digital Desk: Shared editing of paper documents, Technical Report EPC-93-108 Xerox Research Centre Cambridge Laboratory, Cambridge UK., 1993