

Virtual Presence for Medical Procedures

by  
Federico Menozzi

Senior Honors Thesis  
Department of Computer Science  
University of North Carolina at Chapel Hill  
April 27, 2016

# **Abstract**

As medical training becomes more and more complex, with students being expected to learn increasingly specialized and sophisticated procedures, the current practice of having students physically observe all procedures is becoming increasingly difficult. Some procedures are exceedingly rare, while others may rely on specialized equipment not available to the student's institution. Additionally, some procedures can be fast-paced, and critical details might be overlooked in such a hectic environment. We present an application solution that records the procedure with multiple cameras, reconstructs the 3D environment and people frame-by-frame, then utilizes virtual reality (VR) to allow the student to walk through the reconstruction of the procedure through time. We also include several post-reconstruction enhancements, such as video playback controls, scene annotations, and introducing new 3D models into the environment. While we present our solution in the context of medical training, our system is general enough to be applicable in a wide variety of training scenarios.

## **I. Introduction**

### **Importance of Medical Training**

Properly training the next generation of doctors and medical professionals is of paramount importance, especially considering the explosive growth of modern medical advances in recent years. Students today are expected to be proficient in performing increasingly complex medical procedures. In addition to medical students, seasoned doctors and other medical personnel can benefit from continuing to learn new procedures and techniques throughout their careers. It is therefore critical that teaching methodologies for medicine continue to improve and strive to be easy, intuitive, and effective.

### **Qualities of Effective Medical Training**

In order for medical training to be effective, it must easily allow for the student to both quickly absorb the knowledge in question and to retain it effectively for professional use afterwards. At the same time, students must be able to consult materials as a reference in the event that they forget subtle details. The current practice of having students be physically present to observe a procedure is an excellent way for them to see first-hand how medicine is practiced in the real world. However, should the student desire to revisit a particular aspect of a procedure, they are forced to either schedule time to observe it physically, or consult less interactive media, such as a textbook or a pre-recorded video. Ideally, the student should be able to learn procedures in an engaging and interactive manner while retaining the convenience and accessibility of more conventional reference materials.

Additionally, the barrier to entry for the student aide material should be low. The student should not have to be extraordinarily technologically-savvy in order to benefit from enhanced forms of medical training, nor should they have to spend precious time becoming proficient in highly-specialized software or systems. The student should therefore be presented with a familiar and intuitive interface for whatever system he or she uses.

Finally, the technology should be easily adopted by hospitals and other medical institutions. In order to facilitate this, it should be easy to distribute to organizations worldwide. Furthermore, the system should be easily extensible and applicable to a wide variety of medical procedures that may be useful to students.

## **Our Contribution**

We present an application that utilizes immersive VR to allow a student to explore 3D reconstructions of procedure through time in an interactive way. We introduce a temporal element, allowing students to walk around and view each step from any angle they choose, which is not possible with media like images or video.

In order to increase interactivity and utility as a learning device, our system comes equipped with several additional features. These include the addition of familiar video playback controls, annotations for important items in the scene, and features not available in current methodologies such as textbooks and pre-recorded video, such as the ability to introduce 3D anatomical models directly into the scene for further review. These features combine to produce a unique experience in medical education that combines the immersion and interactivity of real presence with the convenience and clarity of annotated video playback.

Our system meets all of the requirements of immersion, interactivity, ease-of-use, distributability, and extensibility. As a program written using the widely-used Unity game engine, it has relatively modest requirements to run, requiring only the software engine and the geometric mesh sequences of the procedure to be installed on the host computer. Coupled with an Oculus Rift DK2 headset, our solution is inexpensive and thus available to a wide variety of medical institutions. Furthermore, as a software solution, the system is easily distributable to any organization with access to the Internet. In fact, while we have chosen specific aspects of our system to use (Unity game engine, Oculus Rift headset), our system should, with minimal adjustments, run on multiple platforms and headsets.

## **Case Study: Prostate Biopsy**

In order to evaluate the effectiveness of the techniques that we have explored, we decided the first scenario to be prostate biopsies. There are several reasons why we have chosen to do so. For one thing, prostate biopsies are performed very frequently [1], so we are better able to collaborate with physicians and schedule procedure recording times. In fact, we have, on several occasions, collaborated with Dr. Eric Wallen and others at the UNC Hospital to record live prostate biopsies for later reconstruction.

In addition to their frequency, prostate biopsies are short and simple—they usually last no more than thirty minutes, and typically require only a single physician and a nurse. This simplicity will play an important role later on when dealing with camera placements, and having a procedure whose motions are simple and easily predictable will allow us to attain higher-quality reconstructions.

## II. Prior Work

There is a wealth of prior work in both temporal 3D reconstruction and introducing VR in medical training. While we draw heavily from many such sources, our system aims to improve upon the current state of temporal 3D reconstruction, as well as adding features and capabilities not present in current VR medical training suites.

### Temporal 3D Reconstruction

Perhaps the most significant and relevant recent prior work in the field of 3D reconstruction is Mingsong Dou and Henry Fuchs' paper entitled "Temporally Enhanced 3D Capture of Room-sized Dynamic Scenes with Commodity Depth Cameras" [2]. Their work centers around using Microsoft Kinect cameras to capture scenes as they unfold in real-time, incorporating technologies like Kinect Fusion [3] to track and fuse moving objects. Utilizing both static and dynamic scene capture methods, their method generates meshes from point clouds obtained from color and depth images, which can then be "played back" to provide the illusion of real-time presence. We have drawn heavily from their work to form the foundation of our system. However, in an effort to improve upon the experience for a potential user of our system, we have improved upon their work in several aspects, from improved texturing to optimal initial camera placements. Figure 1 shows results of [2].

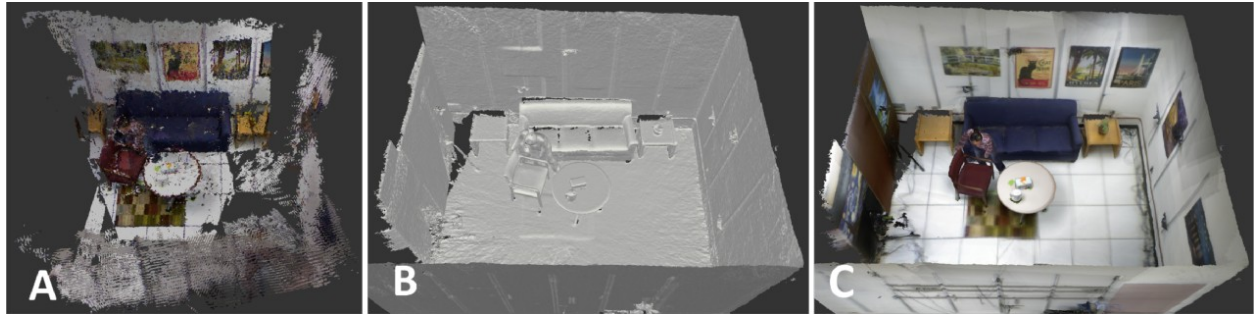


Figure 1: Results of [2]. (A) Original point cloud captured from 10 Kinects; (B) Result of system; (C) Result of system with textures applied. [2]

Other work in this area includes work done at Carnegie Mellon University's Robotics Institute in massively multiview systems for social motion capture [4]. Here, researchers constructed a small room-sized dome of 480 synchronized cameras in order to capture multiple people engaged in social activities. They ultimately used the synchronized video feeds to produce a time-varying 3D structure of anatomical landmarks for each individual. Unlike [2], whose tracking is limited to only one or two people at a time, [4] can handle six or more people interacting with each other at a time. However, because [4] utilizes 480 cameras, rather than the ten or so that we use, they are unable to fit in an ordinary-sized procedure room. Figure 2 shows a video frame captured by the dome and the corresponding skeletal representation of that frame.

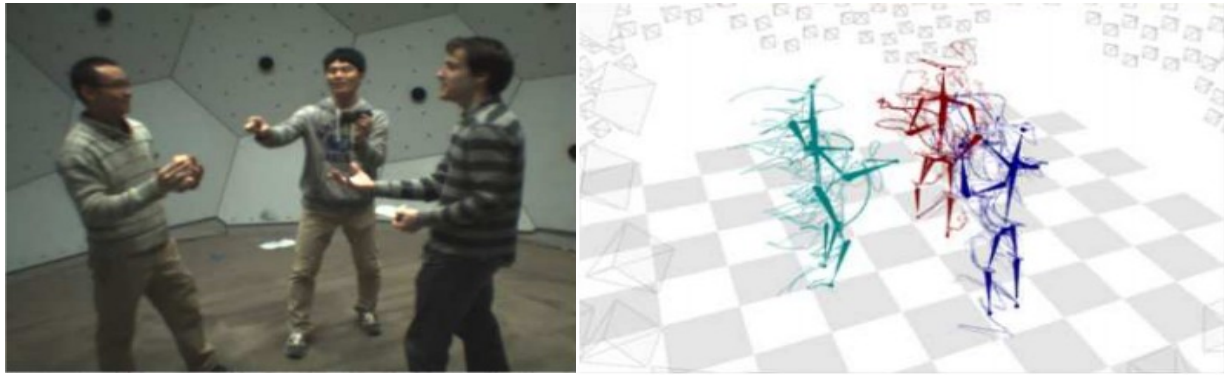


Figure 2: Comparison of video frame (left) with reconstructed skeletal representations (right). [4]

## VR Medical Training

Researchers have long recognized that VR has widespread potential applications to medical practice and training. Studies [5][6] have shown that augmenting surgical training with VR results in procedures with faster completion times and fewer errors.

3D anatomical models have also been shown to be useful in training medical students. In one study [7], researchers constructed a fully-interactive model of the middle and inner ear from magnetic resonance imaging (MRI) scans of a human ear, as seen in Figure 3. In order to test the effectiveness of the model as an educational utility, the researchers created a Web-based tutorial on ear anatomy which they then administered to two groups of students—28 students who had additional access to the ear model, and a control group of 29 students without access to the model.

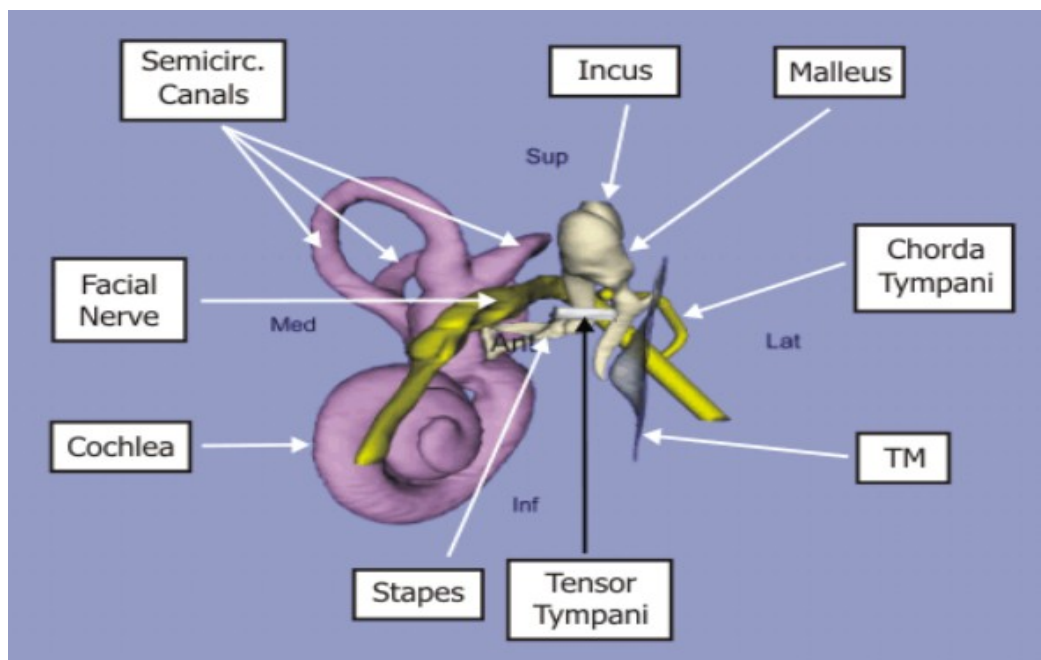


Figure 3. Structures included in the model. [7]

At the end of the tutorials, both groups of students were asked to complete a quiz evaluating their understanding of 3D structures within the ear. The students who had access to the ear model had an average score of 83%, while the control group had an average score of 65%. This statistically-significant difference further demonstrates that medical students with access to interactive media will consistently demonstrate a higher level of understanding of human anatomical structures than those who rely solely on traditional resources.

VR has also been used specifically to enhance trauma surgeries. The 2002 paper “Immersive Electronic Books for Teaching Medical Procedures” [8] explored techniques for utilizing “immersive virtual reality books (ivrBooks)” to combine 2D and 3D data (in both a head-mounted display and wall-based projections) to enhance, among other things, trauma surgery training. Figure 4 shows the ivrBook in action during a liver operation, including a reconstruction, a timeline of the procedure, and a heartbeat monitor, with Much like [2] and our own system, [8] utilizes geometric reconstructions of previously-captured procedure footage that can be played back like video through the use of the ivrBook.

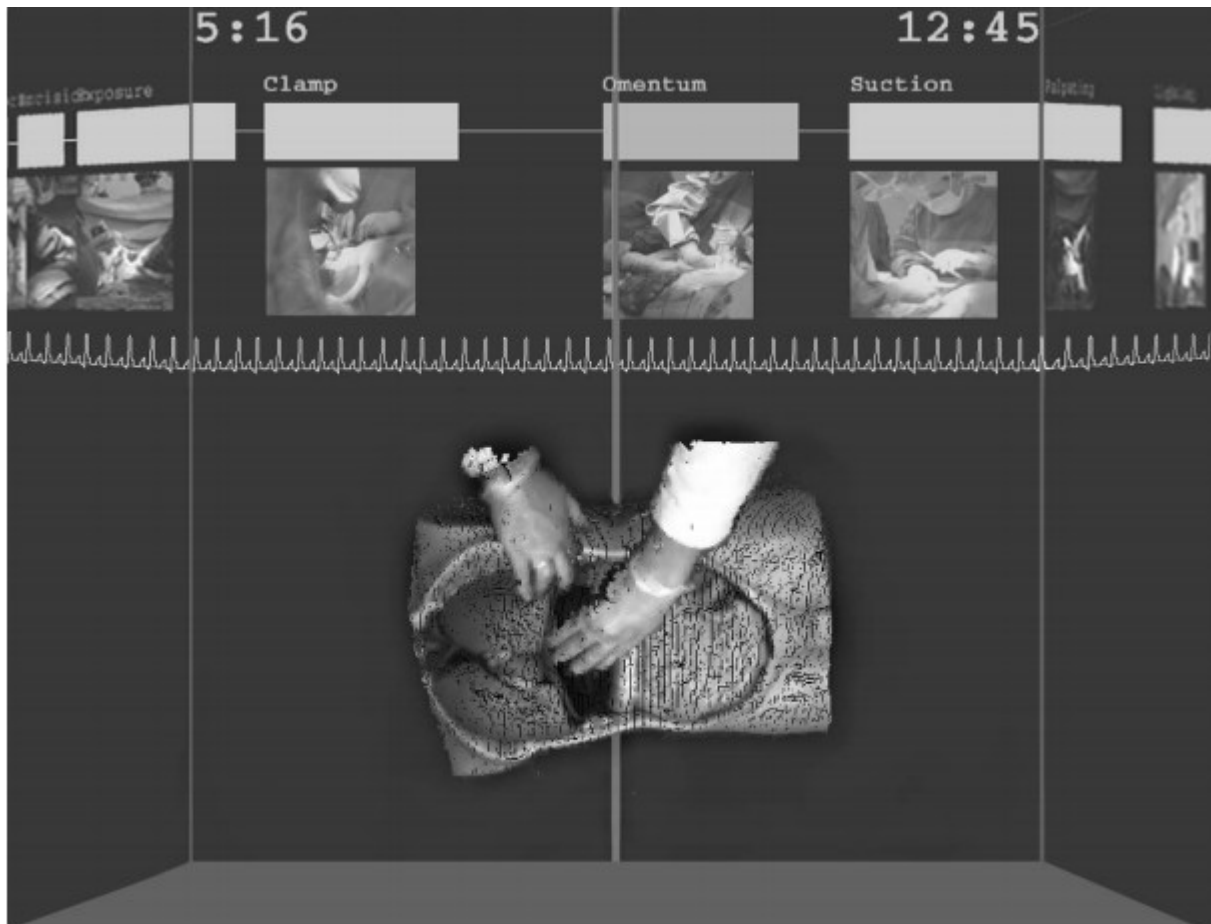


Figure 4: A view of the electronic book in action. [8]

### III. System Pipeline

Our pipeline can roughly be divided into two parts: the temporal 3D reconstruction phase, and the post-reconstruction enhancements phase. The reconstruction phase, the result of work by graduate students Young-Woon Cha and Rohan Chabra, is based on [2], but features several improvements to the pipeline, so it is necessary to discuss this phase in depth. The post-reconstruction enhancements phase is completely new and provides features to make the system useful to medical students. This includes playback controls, procedure annotations, and the ability to introduce new 3D anatomical models directly into the scene.

#### Temporal 3D Reconstruction Phase

The temporal 3D reconstruction phase is composed of three parts: the static reconstruction, the dynamic reconstruction, and the static-dynamic alignment. We will describe these parts in detail, taking particular note of the features that were added to improve upon [2].

##### Static Capture

Before the procedure begins, a single hand-held Kinect is used to slowly sweep the room and record a sequence of color and depth images that will later be used to construct a single 3D mesh of the static scene. Here, “static scene” refers to all of the objects in the scene that do not deform and either do not move at all or move a minimal amount. Unlike [2], at this time we do not make a distinction between “static” (e.g. walls, floor, ceiling, large furniture) and “semi-static” (e.g. chairs, small tables, trays) objects.

It is important that we not rely entirely on dynamic capture data to provide background data because the point cloud from the dynamic capture is much more susceptible to mitigating factors like noise and occlusion. In fact, there are several advantages to using a single, moving Kinect to capture the static scene. First, we are able to keep the Kinect at the optimal distance (roughly one meter) away from the surfaces in order to minimize depth errors. Second, we avoid the laser interference that results from multiple Kinects attempting to estimate depth patterns near each other. Figure 5 shows the results of such interference in the depth image of a Kinect camera. Finally, because the Kinect is not fixed in space like in the dynamic capture, we are able to acquire multiple views of the same geometry, thus increasing the accuracy of the resulting mesh.

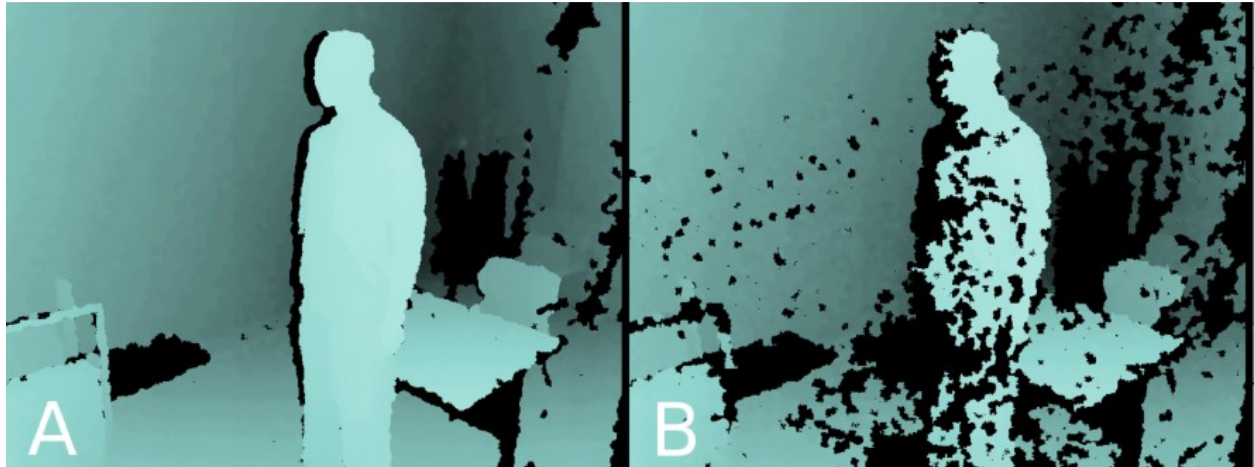


Figure 5: Effects of interference in Kinect depth sensors: (A) Depth map of room with a single Kinect. (B) Depth map of room with multiple overlapping Kinects, producing interference pattern. [9]

After the sequence is captured, the frames are run through a pairwise-matching algorithm in which features from each frame are mapped to corresponding features in other frames. These features allow for the scene to be reconstructed into a single 3D mesh, similar to Figure 6.



Figure 6: Sample static reconstruction of mock procedure room.



## Optimal Camera Placement for Dynamic Capture

One novel feature of our dynamic capture system is the method in which we initially orient the cameras along the walls. Rohan Chabra started a system in which the location and orientation of each of the cameras are determined systematically, rather than relying on human judgment. His system utilizes a simulation to systematically search the space of feasible camera poses, optimizing for minimal occlusion and interference and maximal coverage of important areas, such as the principle surgical site. In addition, the system utilizes basic animation to estimate the motions that participants in the recorded surgical procedure make. Figure 7 shows the results of an experiment Chabra ran with a simple 3-Kinect setup, highlighting the increase in quality with the new camera placements.

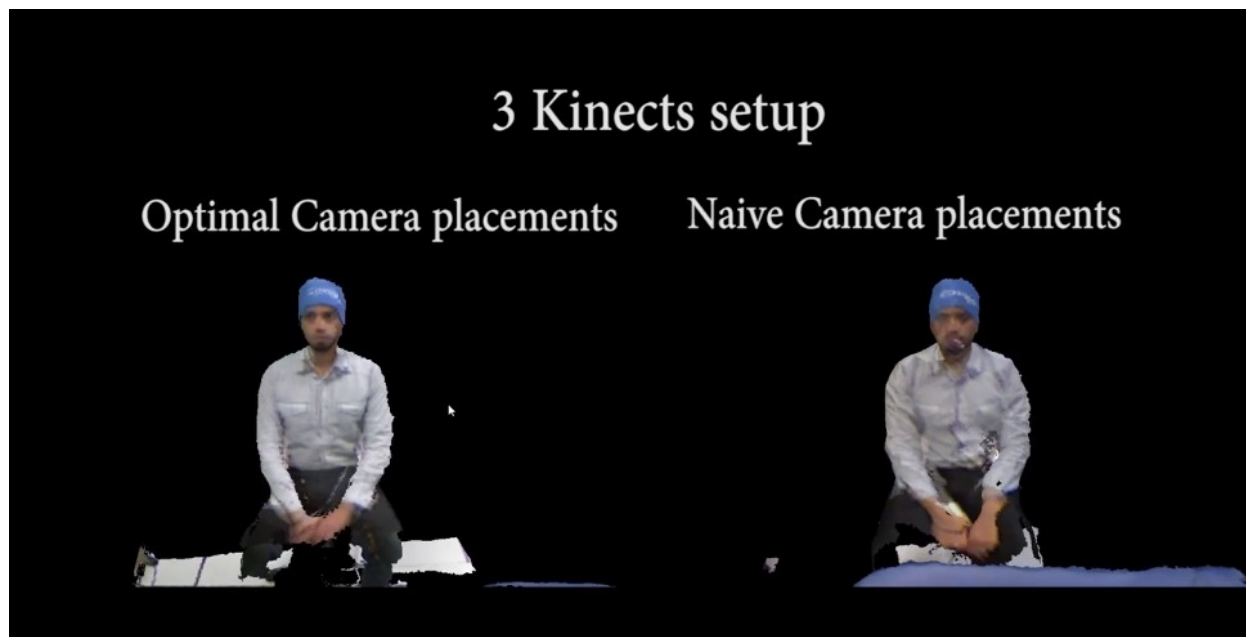


Figure 7: Comparison of naïve vs. optimal camera placements. Notice the increased resolution and detail of the face.

This is one aspect of our pipeline that differs from [2] significantly—[2] uses “naïve” camera placements, determined using only human judgment.

## Dynamic Capture

After the static prescan, a constellation of Kinects is set up in the procedure room for the dynamic capture of the procedure as it unfolds, similar to Figure 8. After the procedure is completed, each camera will have a sequence of color and depth images. If there are  $n$  cameras that each capture a sequence of  $m$  color-depth image pairs, then there will be  $m$  total dynamic reconstructions, each created from the  $n$  color-depth image pairs from the cameras at frame  $0 \leq i \leq m$ . If we “play back” these reconstruction in a VR headset at a framerate similar to that of the Kinect cameras used to capture the procedure, we can achieve the illusion of physical presence, similar to how playing back frames captured by a video camera at a high enough framerate creates the illusion of moving objects. Currently, our system is able to capture data at around 20 frames per second. Additionally, before the frame is fused with the background, the residual

background data (Figure 9) is automatically segmented out (Figure 10).

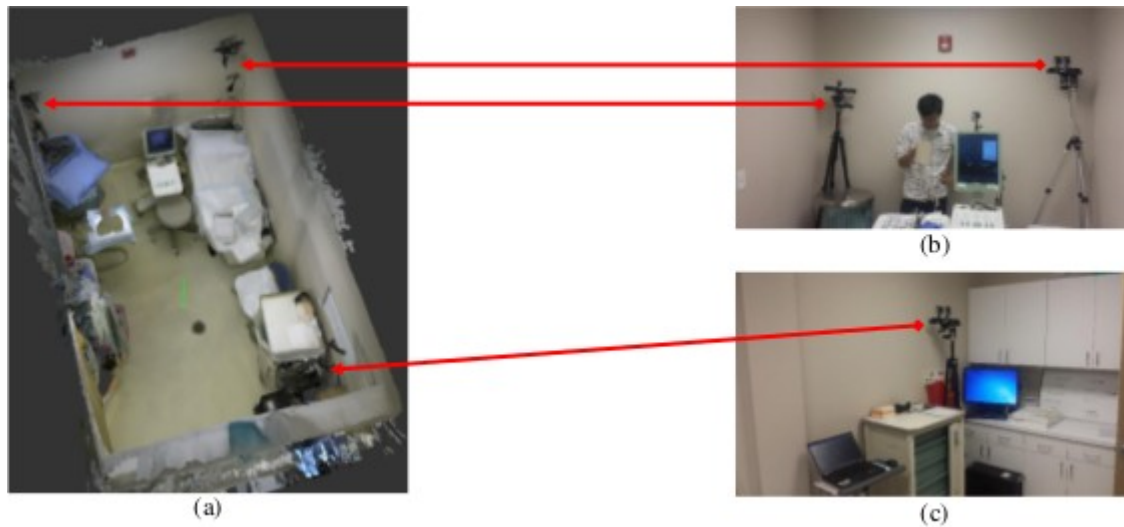


Figure 8: Arrows indicate spatial relationship between mounted Kinects in (a) with those in (b) and (c) for a sample procedure. [10]

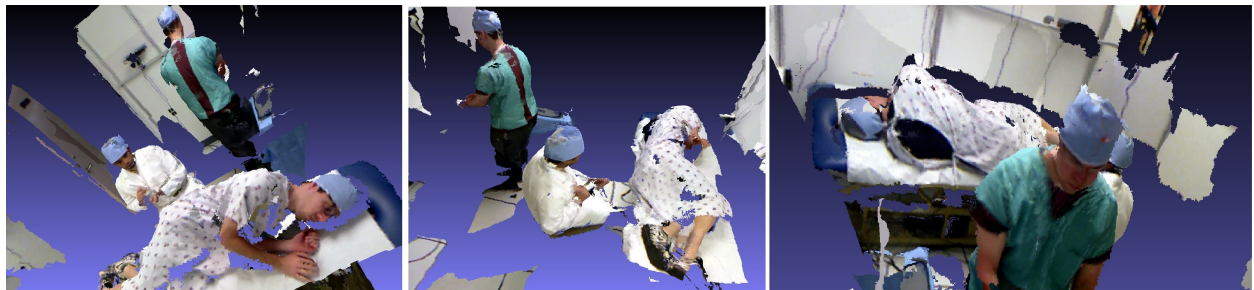


Figure 9: Three different views of the same 3D reconstruction of a single frame before residual background data is segmented out.

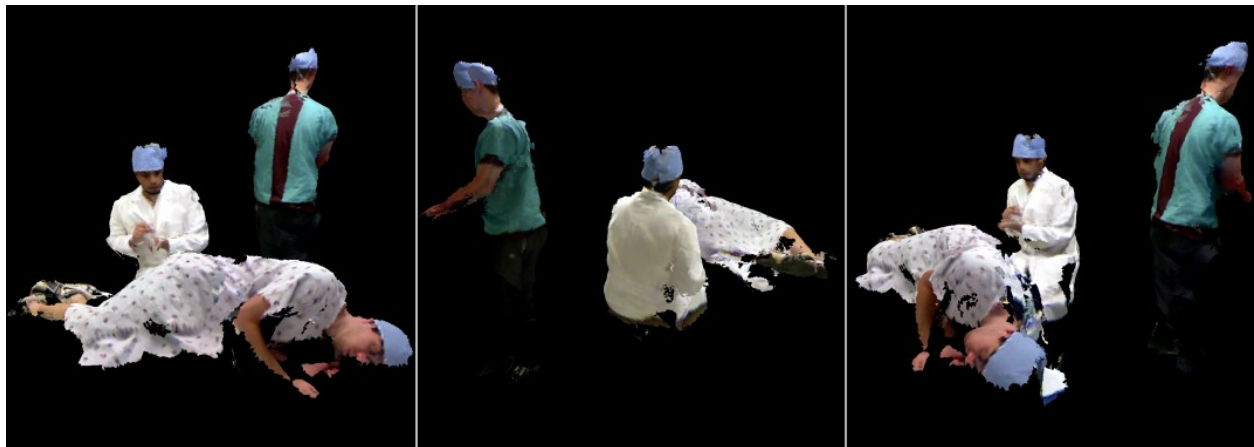


Figure 10: Three different views of the same 3D reconstruction of a single frame after residual background data is segmented out.

## Alignment

Once both the static and dynamic sequences have been captured, the results must be fused into a single combined dynamic sequence with a feature-based alignment technique. During this step, the floor of the static scene is aligned with the floor in the dynamic sequence. In each frame, the two scenes are fused together, creating a single resulting sequence that can later be played back like traditional video. Figure 11 shows the state of the dynamic and static reconstruction at each step of this process.

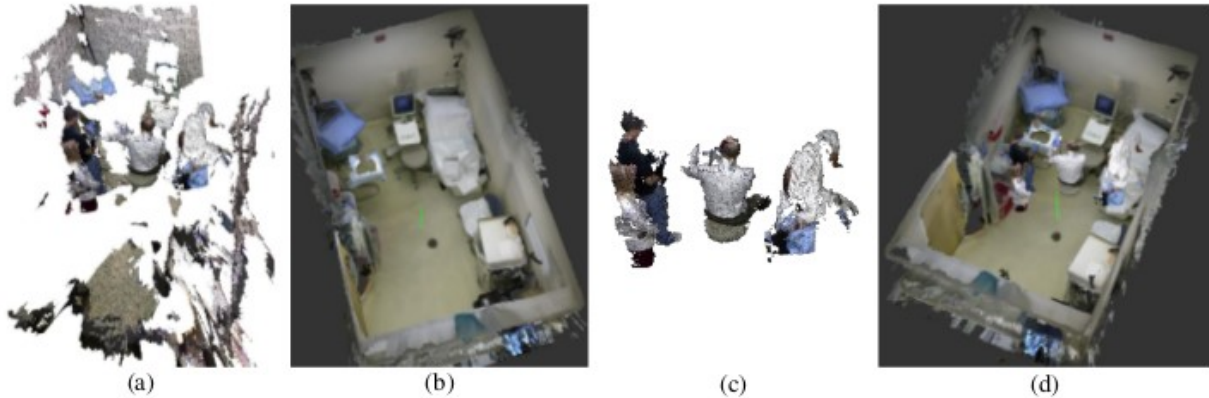


Figure 11: Alignment process: (a) Raw dynamic reconstruction. (b) Static scene reconstruction. (c) Dynamic reconstruction with background data automatically segmented out. (d) Alignment of static and dynamic reconstructions. [10]

## Post-Reconstruction Enhancements

The post-reconstruction enhancements phase takes the results of the temporal reconstruction phase (i.e. the dynamic sequence aligned and fused with the static scene) and adds several features that may be useful for medical training purposes. These features become especially effective when viewed through a VR headset, as they are especially designed to increase learning in a virtual environment.

### Playback Control

Perhaps the most important added feature is the ability to “play back” the frames in 3D with familiar, VCR-like controls. This temporal manipulation allows the user the ability to pause, rewind, fast-forward, etc. through the scene as they see fit, providing a high degree of control in their learning process. Students can toggle the control panel on and off, allowing them to remove the panel from the scene if they find it too obstructing.

### Scene Annotations

We have also added the ability to add customizable annotations to the scene. These can be useful for doctors or other supervising medical personnel to add training instructions or reminders for critical portions of a procedure. The annotations appear as yellow text in the scene. For our

prototype, we place them along one of the walls, easily visible to the user, but they can be placed anywhere, allowing for custom locations for different scene situations. Figure 12 showcases scene annotations and the playback panel in a sample scene.



Figure 12: Screenshot of mock procedure playback viewed in an Oculus Rift headset and rendered with the Unity Game Engine. Notice the annotation along the wall in the scene and the playback control panel floating in front of the user's face.

### Introduction of 3D models

Additionally, our system allows users to introduce 3D anatomical models directly into the scene. This is useful for students who need to reference relevant anatomy for a procedure that they are observing. Like the playback control panel, the models can be switched on and off. Figure 13 shows the scene before and after a new 3D model is introduced.

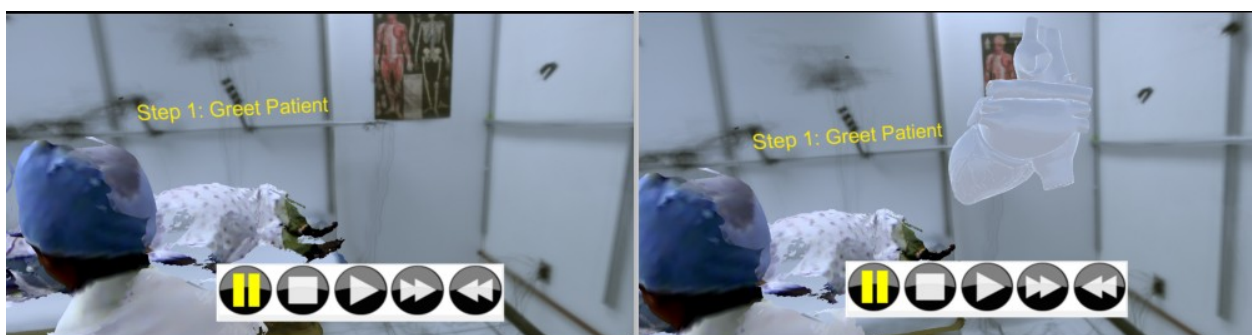


Figure 13: Comparison of scene before (left) and after (right) introducing a new 3D model.

## IV. Advantages

We believe that our system provides several potential advantages when compared to the current

state of the art, both in the temporal 3D reconstruction phase and the post-reconstruction enhancements phase for use in medical training.

### **Benefits of Physical Presence**

The main goal of this system is to emulate physical presence as much as possible while also providing features that would be difficult or impossible to incorporate into a live procedure. Our hope is to provide the same level of immersion and interactivity that one would have in a live procedure, even in the presence of enhancements like frame playback and scene annotations.

### **Accessible**

Despite utilizing a technology as new as modern VR, our system maintains a low barrier to entry for medical students. The wireless mouse is a familiar device to most young people as is therefore well-suited for our initial prototype controller. Additionally, the added features take advantage of tried and true user interface techniques to make adoption of the software as fluid as possible. One example of this is the VCR-like playback controls for advancing through frames. Because most people are familiar with such playback controls, we hope that they will be able to jump in without much assistance.

In addition to being easy to use, our system is accessible from a convenience standpoint. One of the biggest drawbacks to physical presence in a medical procedure is inconvenience, as some procedures are very rare, while others might use equipment or facilities not available at the student's institution. Our system takes advantage of the principle benefit of VR: the ability to be transported to any environment from the comfort of your living room.

### **Extensible**

While our early prototype was designed around a prostate biopsy, our system is generic enough to apply to a wide variety of procedures. The temporal reconstruction can be applied to any indoor dynamic scene, as well as the wall annotations and playback controls. In fact, the reconstruction data can almost be seen as an input to the application as a whole. As such, extending the application to run with different procedures is primarily a matter of using different underlying reconstruction data.

### **Distributable**

Because our system relies on hardware that a hospital would either likely already have at hand or else could relatively easily acquire, only the software and accompanying data needs to be distributed. Therefore, obtaining the system data is as simple as an Internet download, and new procedures could just as easily be downloaded when they become available.

## **V. Limitations**

Our system also has a couple of limitations that principally stem from the fact that we are currently only considering prostate biopsies for our case study. We hope to expand our system in

the future to include more general procedures, but until then, several aspects of our system will not be as optimal as they could be.

### **Only Prostate Biopsies**

Our decision to use prostate biopsies as a case study was helpful in prototyping this system. The ubiquity and frequency of this procedure ensured that we could collaborate often with the UNC Medical School in recording procedures. Additionally, our knowledge that the procedure was a prostate biopsy, combined with our observations of actual prostate biopsies being performed, allowed us to tailor Chabra's simulation to account for specific motions made during typical prostate biopsies. This means that the camera poses that the simulation produced were tailored for this specific procedure, and would thus unlikely work as well for other procedures. Therefore, as it currently stands, our system can only effectively handle prostate biopsies, as a different procedure would require us to use a different simulation for the optimal camera poses. However, this should be easy to change for another procedure, and all other aspects of our system should work unaltered for other procedures.

### **Unity**

In order to hit the ground running and facilitate rapid prototyping, we utilized the Unity game engine to render the reconstructions and incorporate the various post-reconstruction enhancements. Unity abstracts many details about the underlying rendering from the user, allowing them to focus on making 3D games more effectively and with less headache. Unfortunately, while this made the overall process of enhancing the scene more straightforward, we found that Unity sacrifices control and, in some cases, performance, to achieve this goal of user-friendliness. Indeed, the lack of easy multithreading proved to be painful when attempting to create a background mesh-loader. Additionally, we sometimes found that the framerate during dynamic playback was highly variable, ranging anywhere from 50 to 70 frames per second. This behavior was not observed using graphics APIs that give more control the the rendering pipeline, such as OpenGL. Because of this, a future iteration of this system will likely utilize OpenGL to provide better control with multithreading and consistent framerate.

### **“Hard-Coded” Values**

Certain aspects of the post-reconstruction enhancements, such as the poses of the annotations and 3D model, are “hard-coded” into the system. This is because we the programmers have a priori knowledge of the coordinate system of the specific procedure we're dealing with, so we can ensure that the annotations are placed along the wall, for example. Because of this, if we attempted to use a different set of 3D reconstruction data in our current system, the annotations and model would likely not be positioned in a way that makes sense to the user.

## **VI. Future Work**

While our system provides a first step for exploring virtual worlds in a medical context, there are several aspects that could be added or improved upon to further enhance medical training in a variety of ways.



## **Automatic Segmentation**

One nice feature of [2] is that certain “semi-static” objects (i.e. objects that do not deform but occasionally move, such as small furniture) can be manually segmented out from the static reconstruction and tracked during the dynamic sequence, all while retaining the high quality of the fusion and tracking used in the static reconstruction. However, these segmentations must be done manually, after the static reconstruction has concluded, which can be tedious. Graduate student Young-Woon Cha is currently working on techniques to allow for automatic segmentation of “semi-static” objects, increasing the quality of such reconstructions during playback.

## **Ease Of Use**

Ease of use is an important attribute of our system, and ensuring a smooth experience for medical students and physicians is of paramount importance. There are a few aspects of our system that could see improvements in this area. One such example is the placement of the annotations and models in the scene. Currently, the position and orientation of each annotation and model is determined beforehand by developers familiar with Unity and the coordinate system of the scene. There is therefore no easy way for a doctor to adjust these parameters during playback if he or she determines that they would be better suited positioned elsewhere. One possible solution to this would be to have the physician run through a small “configuration utility” whereby they use the headset to walk through the scene and select points in space to position annotations and models. These positions would then be saved to disk to be used in subsequent playbacks by students. Should the physician change their mind, they would be able to run the configuration utility again to make changes.

Another important aspect is the control interface for the student during playback. Currently, the student uses a wireless mouse to control playback and introduce models into the scene. They are therefore limited to only a few possible controls—left and right clicking, middle clicking, and scrolling. It would be beneficial to introduce a more flexible yet easy-to-use controller that is more suited to interacting in virtual worlds. Such a controller would also allow for more actions than a traditional mouse, which could manifest as possible new features in future iterations of the system. For example, our system currently only supports toggling one reference anatomical model at a time with the middle click button, a decision made to show the possibility of the concept while staying within the limits of our controller options. A more sophisticated controller could, for example, allow for more models to be used, or allow the student to manipulate the anatomy directly.

## **Multiple Participants**

Our system was designed with a single participant in mind—the medical student. It's possible for multiple additional participants to wear VR headsets and watch the same dynamic scene unfold as the principle user. However, there are a number of questions that arise in such a situation: who is responsible for controlling playback? What do users see when they look in the direction of another participant?

If we choose to continue utilizing VR, then answering these questions can be difficult, especially the question of what users should see when looking in the direction of other users. Because we must construct our entire environment in VR applications, we would have to render some, if not all, of the other user's body. This can be difficult to do well without breaking immersion, and would necessitate the use of external sensors to track body movements, which may prove too cumbersome to be useful.

Alternatively, we could consider augmented reality (AR) instead of VR. AR has the benefit of not requiring us to have to render much of the surrounding environment (including other users), allowing us to focus on rendering the reconstructions of the dynamic portion of the procedure. Additionally, AR would better facilitate student-teacher interaction, as the instructor could choose to view the scene through his own AR headset. However, AR comes with its own set of challenges: we would no longer have control of how we render the static background content, so the viewing area for an AR procedure would have to be somewhat similar to the room in which the actual procedure took place in order to be believable. In addition, the options for commodity AR hardware are quite limited—most consumer solutions have a much narrower field of view than our Oculus Rift headset, usually about 40 degrees compared to our 100 degrees.

### **Further Scene Enhancement**

Playback controls and scene annotations are not the only ways to add useful information to a system such as ours. There are several additional features that could be added to improve utility, immersion, etc. Such enhancements could be additional features that help to better replicate the live environment or augmentations to the scene not available in live procedures or with current methodologies.

For example, many medical procedures require the use of some sort of live imaging equipment, such as ultrasound. In such a procedure, it would be useful for students who observe the imaging portion of a procedure to be able to view the results of such imaging on the equipment monitor in real-time. The image frames would be responsive to changes in playback so that they pause, play, fast-forward, etc. in tandem with the scene.

There are also additional features not available with traditional medical training technologies that could be added to our system. One key enhancement is the ability to make certain portions of the environment, such as the patient's anatomy, transparent during portions of the procedure. For example, we use a prostate biopsy as our example procedure to prototype our system. During this procedure, the supervising physician may want to point out the proper probe insertion angle to an onlooking student. However, it is currently difficult for the student to get a good view of the insertion due to the rest of the patient's anatomy acting as an obstruction. If it were possible to make portions of the anatomy transparent, the student could get an unobstructed view of crucial details of the procedure, greatly enhancing their understanding of the procedure in ways not possible with current technologies.

However, implementing such a feature would be difficult for a number of reasons. For one thing, the user interface for such a system would not be trivial—how does one easily specify arbitrary portions of the meshes in the environment to cut away? There is also the question of what to



render when a portion of a mesh is cut away—because the meshes are created and textured from 3D reconstructions, there is obviously no information about the inside of a mesh. One possible way of filling in this information would be to track and deform a model of a standard human body, so that at any point in time the model can act as the underlying anatomical structures if the main mesh is cut away.

## Acknowledgements

This research was supported in part by Cisco Systems, by Microsoft Research, by NSF Award II-1405847 (“Seeing the Future: Ubiquitous Computing in EyeGlasses”), and by the BeingThere Centre, a collaboration between ETH Zurich, NTU Singapore, and UNC Chapel Hill, supported by ETH, NTU, UNC, and the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative.

We would like to thank Christopher Paterno for assisting with UNC Medical Institutional Review Board approval procedures and for collaborations on recordings in the clinic, and Jim Mahaney for management of recording equipment for our 3D capture experiments.

## References

- [1] Zlotta, A. R., & Nam, R. K. (2012). “To biopsy or not to biopsy—thou shall think twice.” *European urology*, 61(6), 1115-1117.
- [2] Dou, Mingsong, and Henry Fuchs. "Temporally enhanced 3D capture of room-sized dynamic scenes with commodity depth cameras." *Virtual Reality (VR), 2014 IEEE*. IEEE, 2014.
- [3] Newcombe, Richard A., et al. "KinectFusion: Real-time dense surface mapping and tracking." *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*. IEEE, 2011.
- [4] Joo, Hanbyul, et al. "Panoptic Studio: A Massively Multiview System for Social Motion Capture." *Proceedings of the IEEE International Conference on Computer Vision*. 2015.
- [5] Seymour, Neal E., et al. "Virtual reality training improves operating room performance: results of a randomized, double-blinded study." *Annals of surgery* 236.4 (2002): 458-464.
- [6] Gallagher, A. G., et al. "Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR)." *Endoscopy* 31.4 (1999): 310-313.
- [7] Nicholson, Daren T., et al. "Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model." *Medical education* 40.11 (2006): 1081-1087.

- [8] Van Dam, Andries, Henry Fuchs, and Greg Welch. "Immersive electronic books for teaching surgical procedures." *Telecomm., Teleimmersion, and Telexistence* (2002): 99-132.
- [9] Maimone, Andrew, and Henry Fuchs. "Reducing interference between multiple structured light depth sensors using motion." *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*. IEEE, 2012.
- [10] Cha, Y. W., et al. "Immersive Learning Experiences for Surgical Procedures." *Studies in health technology and informatics* 220 (2016): 55.