

# Head-Mounted Display Integration for Orthopedic Surgery

Zhuokai Zhao\*

Johns Hopkins  
University

Sing Chun Lee†

Johns Hopkins  
University

Javad Fotouhi‡

Johns Hopkins  
University

Long Qian¶

Johns Hopkins  
University

Nassir Navab†

Johns Hopkins  
University

## ABSTRACT

Orthopedic surgery has been very commonly conducted these years, yet it suffers from the lack of efficient harmless guidance system. Current guidance system uses X-rays to only provide the images without any tool-tracking. It starts with acquiring multiple X-ray images from different views to locate the point of entry, under the help of a reference tool. The medical instrument is then invaded and moved inside the patient's body with small displacements. A set of anteroposterior X-ray images are acquired during each small displacement, until the target position is reached. The current workflow is harmful and inefficient. It requires numerous X-ray images for placing wires and screws, which not only harms the patient and surgeon in a direct way, but also increases the probability of potentially damaging the patient's soft tissues and nervous system. This paper comes up with an approach that uses camera and markers for tool tracking and displays the tracking data in 3D using a head mounted display. The presented method is less time-consuming, more efficient and prevents the frequent use of 2D X-rays. It could serve as a road sign indicating that integrating head-mounted display for orthopedic surgery is a path that is worth exploration.

**Keywords:** Mixed / augmented reality

**Index Terms:** H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation]: User Interfaces – Interaction styles; I.3.6 [Computer Graphics]: Methodology and Techniques – Interaction techniques; J.3 [Life and Medical Sciences]

## 1 INTRODUCTION

Using medical imaging as guidance system during the operations has become a common and active research area for more than a century. Starting from 1896, when the X-ray imaging was invented, researchers have been developing many medical imaging systems, including Ultrasonography, Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and Camera Endoscopy. However, while the surgeons are confronted with more forms of imaging data, an intelligent integration and visualization of the acquired data are becoming more necessary. Therefore, using head mounted displays and augmented reality as guidance systems to visualize the imaging data has become very popular in the last decade. F. Sauer et al. [1] developed an augmented reality imaging guidance system in 2001. It enabled the surgeon to see the tumor inside the patient's body in real-time while wearing a video-see-through head mounted display. Joerg Traub et al [2]. created a

navigation interface that combines an augmented reality visualization system, which is based on a stereoscopic head mounted display, with the standard navigation interface. Christoph Bichlmeier et al. [3] introduced a virtual mirror method that navigated the surgeon during spine surgeries.

This work focuses on using augmented reality to guide the surgeon during standard orthopedic surgery. It introduces a solution to track and visualize both the outside and occluded parts of the surgical tool in an optical-see-through head mounted display. The proposed method requires tracking the needle orientation, estimating the needle tip position, interactively classifying the outside and occluded tool segments and augmenting the virtual surgical tool which aligns with the physical tool.

A survey that explores the best visualization cues for surgeons to perceive depth and entry point position information is also included in this work.

## 2 APPROACH

The system setup is shown in figure 1.  $T_{Marker}^{Camera}$  and  $T_{Marker}^{HoloLens}$  denote the tracking results from marker to RGBD camera and HoloLens respectively. Similarly,  $T_{Camera}^{HoloLens}$  represents the calibration result between the two devices, which is the transformation matrix from RGBD camera to HoloLens.  $P_{tip}^{Marker}$  indicates the needle tip position with respect to marker coordinate.

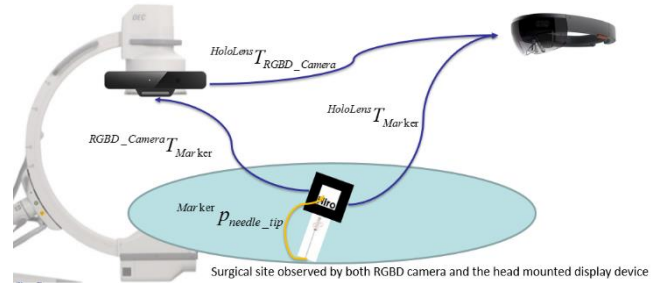


Figure 1: System setup diagram

Since the focus of this project is to augment a virtual surgical tool which aligns with the physical tool, the complete approach can be divided into five main sections: needle orientation tracking; needle tip position estimation; RGBD camera to HoloLens communication; outside/occluded needle segments classification; and virtual needle augmented visualization. The method of using both RGBD camera and HoloLens was designed and developed at the beginning, but was replaced by the idea of using the HoloLens alone in the end.

### 2.1. Tool Tracking with ARToolKit

#### 2.1.1. Tool Tracking using PC and RGBD Camera

In this approach, the needle is tracked by an external PC-connected RGBD camera, through a marker that is rigidly attached to a tool. As showed in figure 2, the pen model on the left is used mostly while developing. It is mainly used for testing tracking system because it is not sharp enough to be inserted into the patient model.

\* e-mail: [zzhao30@jhu.edu](mailto:zzhao30@jhu.edu)

† e-mail: [singchun.lee@jhu.edu](mailto:singchun.lee@jhu.edu)

‡ e-mail: [fotouhi@jhu.edu](mailto:fotouhi@jhu.edu)

¶ e-mail: [lqian8@jhu.edu](mailto:lqian8@jhu.edu)

† e-mail: [nassir.navab@jhu.edu](mailto:nassir.navab@jhu.edu)

The surgical tool model on the right is the real medical tool which can be inserted into the patient's body. It also has the marker that is 50% size of the original marker which is attached to the pen model. Although in theory smaller marker is not as stable as the big marker in terms of tracking robustness, it is in fact an engineering trade-off because the field of view of HoloLens is relatively narrow, which makes the big marker inconvenient during the use. The tracking algorithm is implemented with ARToolKit.

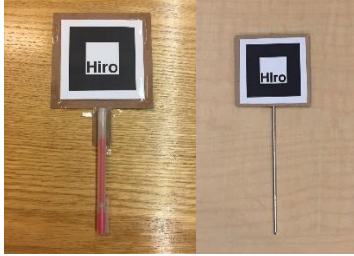


Figure 2: Pen model (left) and surgical tool model (right) with markers used for tracking

The tracking result was formatted to be the 4x4 homogeneous transformation matrix from marker to the RGBD camera, which is denoted as  $T_{Marker}^{Camera}$ .

### 2.1.2. Tool Tracking using HoloLens

With the front-facing 2-megapixel image/video camera (locatable camera) available on HoloLens, which is indicated in figure 3, marker tracking using HoloLens could ideally be very similar to using the RGBD camera with PC. However, since HoloLens runs on Universal Windows Platform (UWP), which uses different tools and run-time libraries than on the PC, ARToolKit does not support any UWP devices. Therefore, the solution implemented above could not be deployed to HoloLens directly. Luckily, a third-party library called HoloLensARToolKit, which contains similar functions to those included in the PC-based ARToolKit, has been published [4] and is ready to use.

Same as before, the tracking result was formatted to be the 4x4 homogeneous transformation matrix from marker to the HoloLens, which is denoted as  $T_{Marker}^{HoloLens}$ .



Figure 3: HoloLens locatable camera

### 2.2. Tool Tip Position Estimation

The needle tip position is estimated through the pivot calibration. A separate program was developed in Unity. It runs in real-time to save all the transformation matrices from marker to camera in a text file. As shown in figure 4, the green 4x4 matrix on the upper right corner indicates the current transformation matrix from marker to the camera.

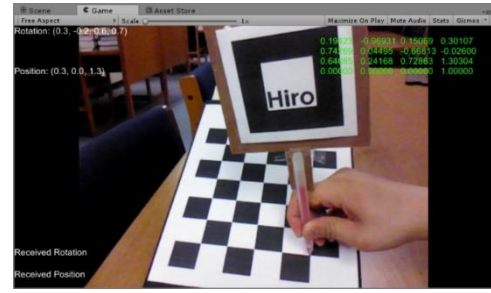


Figure 4: Real-time program to save transformations

The tip position is kept unchanged for the whole time while moving the marker around for transformations of different poses. For each pose, below relationship stands:

$$T_{Marker}^{Camera} \cdot p_{tip}^{Marker} = p_{tip}^{Camera}$$

which is the same as:

$$R_{Marker}^{Camera} \cdot p_{tip}^{Marker} + p_{Marker}^{Camera} = p_{tip}^{Camera}$$

where  $R_{Marker}^{Camera}$  and  $p_{Marker}^{Camera}$  are the rotation and translation components of the homogeneous transformation  $T_{Marker}^{Camera}$ .

Since the surgical tool is rigid,  $p_{tip}^{Marker}$  and  $p_{tip}^{Camera}$  do not change over time. The above equation is further manipulated to be:

$$[R_{Marker}^{Camera} - I] \cdot \begin{bmatrix} p_{tip}^{Marker} \\ p_{tip}^{Camera} \end{bmatrix} = -p_{Marker}^{Camera}$$

By solving the least square solution of the above equation using the transformation results collected, the tip position with respect to the marker coordinate is obtained.

### 2.3. RGBD Camera to HoloLens Communication

To calibrate the RGBD camera and HoloLens, the system needs two requisites: both RGBD camera and HoloLens must observe the marker that is attached to the needle simultaneously; the RGBD camera is able to send the transformation result,  $T_{Marker}^{Camera}$ , to HoloLens in real time.

The first condition is required because it makes the two inputs,  $T_{Marker}^{Camera}$  and  $T_{Marker}^{HoloLens}$  available. The second condition is important because only by that, the transformation from RGBD camera to HoloLens,  $T_{Camera}^{HoloLens}$ , can be calculated within HoloLens as:

$$T_{Camera}^{HoloLens} = T_{Marker}^{HoloLens} \cdot T_{Marker}^{Camera^{-1}}$$

The sending and receiving parts on PC and HoloLens are implemented through the User Datagram Protocol (UDP), which is a communication protocol that is mainly used for establishing low-latency transmissions.

### 2.4. Tool Segments Classification

The first step to classify the outside and occluded parts of the inserted surgical tool is the successful identification of the patient's body surface. Two approaches have been made during the research process. The first one is to use the RGBD camera to generate and segment the point cloud data of the patient's body. This approach was later found to be inefficient during operation because of low transmission rate between RGBD camera and HoloLens via UDP

communication. The second approach takes advantage of the HoloLens's spatial mapping function to detect the surface mesh that is within the HoloLens' field of view. Although only the second approach was chosen as the final implementation method, both approaches are presented and discussed in this section.

#### 2.4.1. Surface Detection with RGBD Camera

To generate the point cloud data that represents the patient's body using RGBD camera, the entire process could be divided into three steps. The first is to generate point cloud data for all the objects that the camera observes; the second is to segment the point cloud data based on surface normal; the third is a further segmentation that uses color.

The RGBD camera has one RGB camera for color captures (RGB image) and one Infrared camera for depth captures (depth image). The relative position and orientation between the two cameras do not change. The algorithm that implements the first step could be presented as:

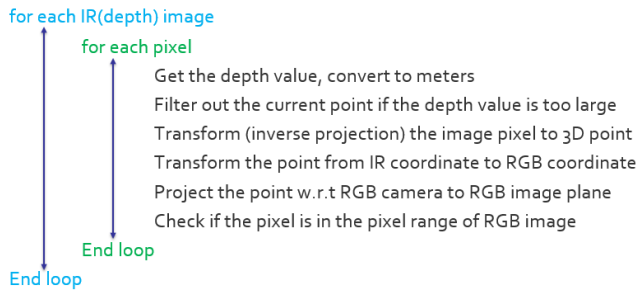


Figure 5: Point cloud construction algorithm

The main idea of constructing the point cloud data in the above algorithm is to find the correspondence between RGB and IR pixels in each video frame to generate the point cloud result. After obtaining the data, segmentation starts with the normal-based region-growing algorithm [5].

The first step of conducting normal-based region-growing algorithm is to generate the normal cloud that corresponds to the point cloud. The algorithm [6] used in this project is concluded in figure 6.

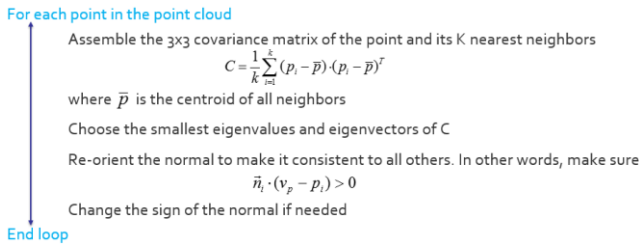


Figure 6: Normal cloud construction algorithm

After both the point cloud and the normal cloud have been determined, the normal-based region growing algorithm [6] could be implemented, with the details concluded in figure 7.

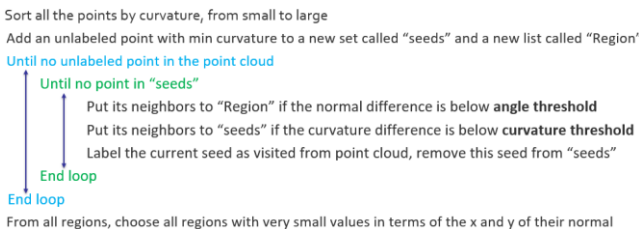


Figure 7: Normal-based region growing algorithm

Since in the scope of this project, the insertion position is on the top surface of the patient body, the algorithm picks the cluster with the normal that is pointing up to be the final segmentation result. However, it is likely that the background environment has surface with the similar normal (floor, etc.) existing, which causes noises in the segmentation result. To further address the issue, the normal-based segmentation result is converted back to unsorted point cloud, and is used for color-based region-growing algorithm [7]. The detail of the algorithm is showed in figure 8.

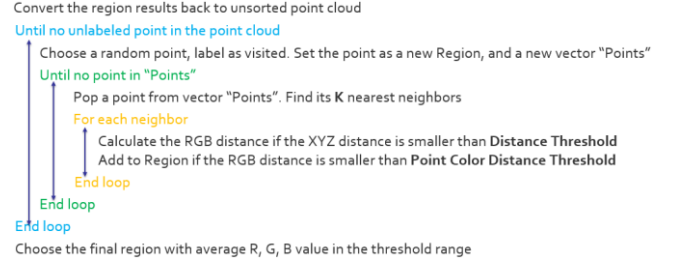


Figure 8: Color-based region growing algorithm

By choosing the appropriate thresholds, color-based region-growing algorithm gives final segmented cluster that satisfies the color requirement of the patient body surface. With the combination of the normal-based segmentation algorithm, the target surface could be successfully segmented even in a relative complex environment, which contains objects that is geometrically very similar to the target object.

The segmentation results are shown in figure 9. The left image indicates the real environment while the right image shows the segmented point cloud result.

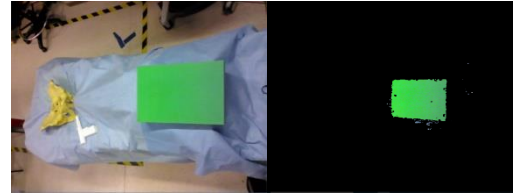


Figure 9: Comparison between real environment (left) and segmented result (right) using RGBD camera

Although the result shown is quite satisfying and stable during the experiment, unfortunately, it was found to be not useful after trying to send the point cloud data over the RGBD camera – HoloLens communication in real-time. The transmission rate as well as the processing ability at HoloLens are not fast enough to handle the point-cloud data in real-time. The issue leads to a more direct and efficient implementation method introduced in section 2.4.2.

#### 2.4.2. Surface Detection with HoloLens

HoloLens can be used to detect the surrounding environment through its spatial mapping ability, which could be implemented to generate virtual surface mesh on the real-world surfaces [8]. More specifically, this is achieved by 4 environmental understanding cameras that are on the HoloLens, where each of the camera has a 512x424 time-of-flight image sensor with multi-frequency photo-demodulation. The 4 cameras are used with the state-of-art Simultaneously Localization and Mapping (SLAM) algorithm that Microsoft developed [9] and produce accurate and robust



surrounding environment reconstruction result. The surface mesh, rendered using white triangular, is shown in figure 10.

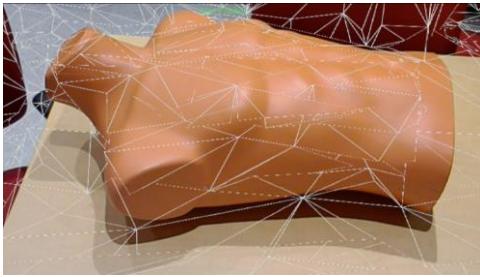


Figure 10: Spatial mapping effect

## 2.5. Virtual Tool Augmented Visualization

### 2.5.1. Augmentation Visualization on PC

The augmentation visualization program runs in Unity on PC, with ARToolKit imported as an external package. In real-time, the program captures the marker's position and orientation, loads the pivot calibration result, and computes the current tool tip position with respect to the camera. The length of the virtual tool is user-defined, and the focus is to align the virtual tool along with the physical tool. The result with the virtual tool augmented in color blue is shown in the figure 11. The first image at the top was augmented on the initial tool model with 3cm virtual tool length. The second image at the bottom was implemented on the real medical tool with a 10cm virtual tool.

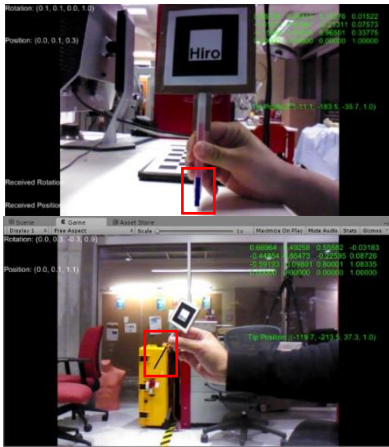


Figure 11: Virtual tool augmentation with PC on pen tool model (top) and surgical tool model (bottom)

### 2.5.2. Augmentation Visualization on HoloLens

As discussed, there are two ways to track the tool, either using PC and RGBD Camera or using HoloLens alone. Therefore, there are also two ways to display the virtual tool within HoloLens. More specifically, the displaying method is the same, but the input data, which is the start and end positions of the virtual tool, could be acquired through different tracking methods.

The first method is to transform the virtual tool's positions from RGBD camera coordinate to positions that are with respect to HoloLens. This approach requires the calibration process between RGBD camera and HoloLens, as described in section 2.3. The advantage of this method is that less computation is required for HoloLens, which improves the smoothness of tool motions because of the relatively weak computational power that HoloLens has. However, the disadvantage is that keeping the marker being visible

to both RGBD camera and HoloLens simultaneously could be hard, especially when the HoloLens's field of view is already narrow. Furthermore, using external camera produces a larger error. This intuitively makes sense, because it combines errors from both the camera tracking and the Camera-HoloLens calibration. This larger error could be directly observed in the first image of Figure 12.

The second method is to use HoloLens alone and perform all the marker tracking and pivot calibration integration within HoloLens. The advantage of this method is that keeping the marker visible to both is no longer required. Surgeons will have more freedom while doing the operation without worrying too much about losing the tracking results. Also, the final augmentation is more accurate, as shown in the second image of Figure 12. The latency is slightly more obvious than the first scheme, but not a significant difference, and should not be a huge concern.

Besides the two methods, which provide the input positions of the virtual tool, there is another step that must be accomplished to get an ideal displaying result in HoloLens. Unlike augmenting on 2D videos/images, augmenting the virtual objects on physical objects using optical see-through head mounted displays (OST-HMD) like HoloLens would require an additional calibration. The aim of such calibration is to compute the transformation from world coordinate to HoloLens's holographic display coordinate so that the virtual objects can be represented in the same coordinate system as the real object. Many researchers have developed many solutions to this problem with respect to accuracy [10, 11], robustness [12, 13, 14], and user friendliness [15, 16]. However, very few of them are aimed for head-mounted displays like HoloLens, which has a holographic display instead of previous monoscopic or stereoscopic displays. The comprehensive approach that is specifically addressing HoloLens's calibration issue does not exist until Ehsan Azimi, et al. [4] published a completed solution which provides a separate calibration process with output as a ready-to-use 4x4 transformation matrix. By applying this paper's method, aligning the virtual tool with the physical tool became successful for both tracking schemes. Two images illustrating results from each scheme is presented in figure 12.

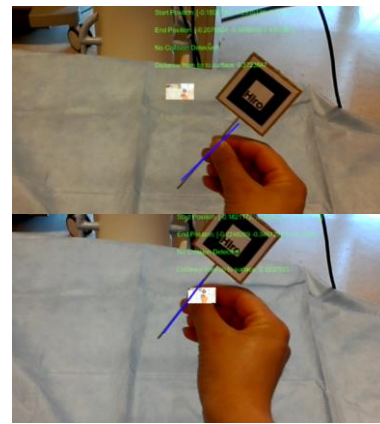


Figure 12: Virtual needle augmentation using external camera (top) and standalone HoloLens (bottom)

### 2.5.3. Other Augmentation Effects on HoloLens

Simple line-model virtual tool augmentation is not enough to guild the surgeon finish the task. Two of the biggest confusing places that surgeons encounter when using head-mounted displays as guidance systems are the insertion position and depth perception. To solve the issue, an invisible laser ray was developed with the starting point same as the needle tip position and the direction same as the needle's pointing direction. The laser ray, although invisible,

collides with the spatial mapping mesh introduced in section 2.4.2. When such a collision happens, a green ring appears at the collision position with the orientation aligning with the surface. The function effect is showed in figure 13. This can help the surgeon to have a much better understanding of where the needle is going to be inserted. It also improves the connection between the real world (patient body) and the virtual object (tool augmentation).

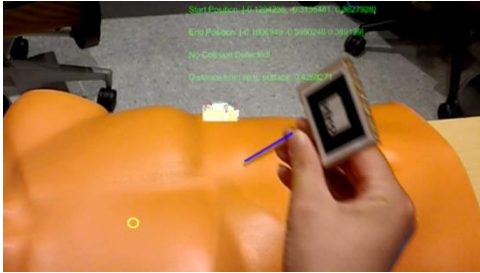


Figure 13: A green ring appears at the position where the virtual tool is going to be inserted

When the tool is inserted into the patient's body, which is detected based on the distance from the entry ring (collision position between the laser ray and surface mesh) to the needle tip, the outside virtual part remains unchanged as it was before the insertion (blue line). However, the inserted part, which is occluded, is rendered in a distinct way. Figure 14 shows a rendering method using different (thinner) widths and gradient color from dark (pink) to light (white) to increase the user's perception on the occluded tool part. This rendering method was chosen based on a user-experience survey that will be discussed later in this paper.

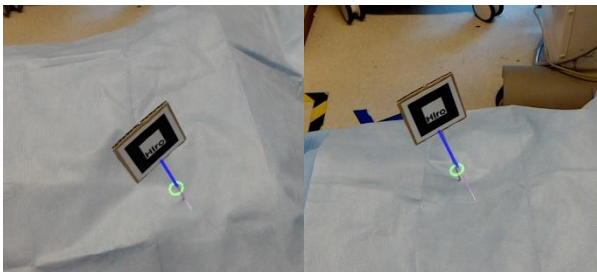


Figure 14: Effects when the needle is inserted

### 3 EVALUATIONS AND DISCUSSIONS

In this section, two major questions are evaluated and discussed. The first one is the accuracy of the virtual tool augmentation on both PC and HoloLens. The second is the survey results obtained from the user-experiment survey, which mainly addresses evaluations and comparisons between different visualization cues implemented to increase the users/surgeons' ability to percept the insertion point on patient's skin and the depth of the inserted part of the surgical tool.

#### 3.1 Accuracy Analysis of Tool Augmentation

The virtual tool augmentation error analysis discussed in this section is the accuracy of the virtual tool tip position with respect to the camera, which could either be the RGBD camera or the camera on HoloLens. More specifically, the error analysis in this section addresses the combined error from both the marker tracking system and the pivot calibration procedure. Although the virtual tool is augmented separately in PC and HoloLens, since the data (tip position) is obtained through the same method, the discussion is not separated in this section.

The experiment system set-up is showed in figure 15. A checkerboard and a camera are positioned where the checkerboard is completely within the camera's field of view. The relative position and orientation between the checkerboard and camera were kept unchanged throughout the experiment. The first step was to obtain the extrinsic of the camera, which is the coordinate transformation from checkerboard to camera. It was achieved by implementing the checkerboard calibration process. Note that the coordinate of checkerboard was defined to be located at the left upper corner of the board. In this experiment, the reprojection error of the calibration process was 0.12 pixels, which is considerably good. The second step was to transfer the corner positions from checkerboard coordinate to camera coordinate by multiplying the extrinsic matrix obtained above with the position vectors. These are set as ground truth of the corner positions with respect to camera coordinate.



Figure 15: Experiment setup for accuracy test

To start the testing process, the tool was placed where the tip was exactly at each corner location of the checkerboard. All the tip positions, obtained from tracking and pivot calibration process, were noted, which were used to be compared with the ground truth determined through the checkerboard calibration.

The error result is shown in table 1. In general, error in all dimensions are around 6 mm. Also, the errors were quite consistent over different sets of locations and experiments. Although orthopedic surgery normally requires 2mm accuracy, considering the relatively large error that ARTToolKit produced, which is around 4-5 mm, the total error is not too bad.

Direction	Error (mm)
x	5.268
y	7.843
z	6.766

Table 1: Virtual tool augmentation error

#### 3.2 User Survey of Perception Cues

As introduced previously, different visualization cues were implemented to help surgeon's perception of depth and entry point position when the tool is inserted into the patient body. Three main different perception cues were used – color, size and entry point indicator. More specifically, color method includes three subcategories – high-contrast color compared to the background color; different colors that represent different tool parts; and gradient-effect color. These three subcategories are shown in the left, middle and right image in figure 16 respectively. Gradient-effect color has three effects: top light (white) bottom dark (pink); top dark (pink) bottom light (white) and a mix of two colors (orange and yellow) from top to bottom. The right image in figure 16 is a demo of the light to dark version.



Figure 16: Different visualization cue effects

The second and third perception cues are the entry point indicator and different widths that were used to represent distinct parts of the needle. Both images in figure 17 include the entry point indicator and the different widths of various parts effect. They were taken from different angles to give a clearer demo.



Figure 17: Different visualization cue effects

The entry point indicator, which is the same as what was shown in figure 13 and 14, is a green ring. It is always on the patient's surface with orientation aligning with the surface. Different widths were simply used to emphasize the part that is inserted by making it thinner than the outside part.

15 people were invited into the survey. Table 2 below shows the basic information about the participants. Most of them are males, and from 18 to 24 years old. About half of them have experienced with AR or orthopedic surgery. Fortunately, none of them experienced discomfort with AR before.

Category	Details	Number	Percentage
Gender	Male	13	86.7%
	Female	2	13.3%
Age	18-24	11	73.3%
	24-30	4	26.7%
Experience	AR	7	46.7%
	VR	2	13.3%
	Orthopedic	7	46.7%
Discomfort	AR	0	0%
	VR	1	6.7%

Table 2: Participants information

Notice that to be regarded as being experienced with AR or VR, the participants need to have at least used the device for many times. On the other hand, being treated as having experienced with orthopedic surgery in this survey is lower-standard – having observed the surgery performed by surgeons is enough.

Table 3 below shows the comparison result between different visualization cues. Note that the definition for better or worse is that the first method listed on the left is better/worse than the second method listed on the right. The comparisons fall into two big categories, which are also the main objectives – entry point identification and depth perception. The entry point perception is visualized through same-color, ring indicator and different-color scheme. From the result, it could be seen that using ring indicator and different-color scheme are both better than simply using the same color. More than half of the participants think that the different-color method is better than using the ring indicator in terms of the entry point perception.

Objective	Rendering Method	Participants Opinion		
		Better	Same	Worse
Entry Point Perception	Ring VS No Ring	100%	0%	0%
	Different Color VS Same Color	93.3%	0%	40%
	Different Color VS Ring	60%	0%	6.7%
Depth Perception	Different Widths VS Same Width	73.3%	0%	26.7%
	Gradient Color (Light to Dark) VS Non-Gradient	38.5%	7.7%	53.8%
	Gradient Color (Dark to Light) VS Non-Gradient	58.3%	0%	41.7%
	Dark to Light VS Light to Dark	41.7%	25%	33.3%
	Mixed Gradient Color VS Non-Gradient	75%	8.3%	16.7%
	Mixed VS Light to Dark	61.5%	15.4%	23.1%
	Mixed VS Dark to Light	33.3%	33.3%	33.3%

Table 3: Comparison between different visualization cues

The depth perception was visualized through using different widths and gradient-color effects, as showed in the table 3 above. The results show that using different widths is helpful. However, not all the gradient-color effects are helpful. Slightly more than half of the total participants believe the depth perception became worse when using the light-to-dark gradient color. However, if it was the dark-to-light or mixed-color, then the depth perception was improved.

Table 4 shows a more detailed result of all the rendering options. There are four basic rendering methods, as introduced above. The “Details” column describes the specific details used in each method. All the methods are evaluated based on four aspects – distinguishing the entry point location, distinguishing the outside and occluded parts of the needle, and the depth perception. The score goes from 0 to 10, as 0 indicating no perception and 10 indicating great perception. The results show that the best option for distinguishing the entry point position and depth perception is the dark-to-light gradient-color method with the ring indicator, distinct colors and widths effects. The most effective rendering method for distinguishing the outside and occluded parts is the light-to-dark gradient-color method. In fact, this introduces a trade-off because the light-to-dark gradient color has a relative poor depth perception compared to other gradient-color schemes. The highest scores are marked as red in the table below.

Rendering Method	Details	Scores			
		Distinguish Entry Point	Distinguish above part	Distinguish occluded part	Depth Perception
Color only	Same Color	5.21	4.73	4.20	4.20
	Two Colors	7.00	7.13	7.13	6.47
Color + Entry Point	Same Color + Entry Point	8.28	7.27	7.13	6.13
	Two Colors + Entry Point	8.57	8.33	8.33	7.93
Above + widths		8.79	8.73	8.73	7.93
Above + gradient colors	Light to Dark	8.92	8.92	8.92	8.23
	Dark to Light	9.08	8.83	8.83	8.83
	Mixed Color Gradient	8.92	8.69	8.84	8.23

Table 4: Average scores of each perception cue

The last set of comparisons is exploring if there is a difference in terms of grading from experienced and non-experienced users. In table 5, the first number in each entry represents the average points from all experienced users, and the second number is the average from non-experienced users. All the blocks are marked green when the experienced users gave more than 0.3 points higher than the non-experienced users, while the red blocks represent the opposite situation, that is, when the non-experienced users gave more than 0.3 points higher than the experienced users. Although it could be directly observed that the non-experienced users tend to give higher scores, to conclude the results, the average points of the three categories: entry point, different parts and depth perception given by experienced users are 7.82, 7.80 and 7.10, while the scores given by non-experienced users are 8.39, 7.72 and 7.25.



Rendering Method	Details	Scores (Experienced / Non-Experienced)			
		Distinguish Entry Point	Distinguish above part	Distinguish occluded part	Depth Perception
Color only	Same Color	5.14 / 5.29	4.57 / 4.86	4.43 / 4	4.29 / 4.13
	Two Colors	6.71 / 7.29	6.86 / 7.38	6.86 / 7.38	6.14 / 6.75
Color + Entry Point	Same Color + Entry Point	7.86 / 8.71	7.29 / 7.25	7.43 / 6.88	6.29 / 6
	Two Colors + Entry Point	8.29 / 8.86	8.14 / 8.5	8.14 / 8.5	7.29 / 8.5
Above + widths	Thinner occluded part	8.79 / 8.43	8.73 / 8.71	8.73 / 8.57	7.93 / 7.71
Above + gradient colors	Light to Dark	8.92 / 8.71	8.92 / 9.00	8.92 / 8.86	8.23 / 8.00
	Dark to Light	9.08 / 8.83	8.83 / 9.17	8.83 / 8.83	8.33 / 8.50
	Mixed Color Gradient	8.92 / 8.57	8.69 / 9.14	8.85 / 9.29	8.23 / 8.57

Table 5: Experienced vs non-experienced users result

#### 4 FUTURE WORK

During the experiment, the biggest issue found for this guidance system is that the field of view of the HoloLens's front-facing camera is relatively narrow. It restricts the ability for surgeons to move the tool freely while having the system track the physical tool stably. Therefore, the future work would be to address this issue. One potential solution for this is to implement HoloLens's spectator view. Spectator view function enables additional HoloLenses as "spectator" to view the virtual display that the host HoloLens has. Microsoft used to have the same issue when displaying the virtual objects in the spectator view due to the narrow field of view it has on the HoloLens. They solved the issue by having an external RGB camera, for example, a Digital Single Lens Reflex (DSLR) camera that has a broad field of view calibrated and replaced for the HoloLens's front-facing camera [17].

Similarly, this idea could be experimented on this project. A HoloLens and the DSLR camera could be the host and mounted to the Mobile C-arm at the position closed to the gantry. In fact, they could be a replacement and upgrade of the camera that is currently mounted on the Camera-Mounted Mobile C-arm (CAMC) system [18]. The system setup would allow the marker tracking system much more robust. The surgeon, while using the system, could have a spectator HoloLens, and by activating the spectator view, to view the virtual needle that is currently implemented in this project within the host HoloLens. Due to the ability that the current CAMC has, it would be also very feasible in the future to integrate the X-ray data and have a mix of the real patient, medical X-ray data, physical and virtual needle display.

Another issue of the current system is that it produces larger error than the error required for orthopedic surgery. As discussed in section 3.1, the main error comes from the current tracking system. Therefore, another future work is to decrease the tracking error that the current tracking system has. This could be possibly done by using multiple-marker tracking or other schemes.

#### 5 CONCLUSION

Existing guidance system for orthopedic surgery is low-efficient, time-consuming and causing more than necessary amount of X-ray dose absorption for patients. It also does not provide an intuitive 3D guidance, which, on the other hand, requires the surgeon to do the complex mental mapping from 2D to 3D themselves. In this present study, an approach that provides a 3D interactive guidance system using HoloLens was developed. Experimental results indicated that although the accuracy of this method does not fulfill the requirements for orthopedic surgery currently, it is very possible to decrease the error by upgrading the tracking system, so that it could be used in medical operations in the future.

In addition, different perception cues that are used to improve surgeon's depth and entry-point position perception ability were successfully experimented and evaluated. A complete survey was designed and conducted.

#### REFERENCES

- [1] F. Sauer, Ali Khamene, B. Bascle, G. J. Rubino, A Head-Mounted Display System for Augmented Reality Image Guidance: Towards Clinical Evaluation for iMRI-guided Neurosurgery, Proceedings of the 4th International Conference on Medical Image Computing and Computer-Assisted Intervention, p.707-716, October 14-17, 2001
- [2] Joerg Traub, Philipp Stefan, Sandro Michael Heining, Tobias Sielhorst, Christian Riquarts, Ekkehard Euler, Nassir Navab, Hybrid navigation interface for orthopedic and trauma surgery, Proceedings of the 9th international conference on Medical Image Computing and Computer-Assisted Intervention, October 01-06, 2006, Copenhagen, Denmark [doi>10.1007/11866565\_46]
- [3] Christoph Bichlmeier, Sandro Michael Heining, Mohammad Rustae, Nassir Navab, Virtually extended surgical drilling device: virtual mirror for navigated spine surgery, Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention, October 29-November 02, 2007, Brisbane, Australia
- [4] Long Qian, Ehsan Azimi, Peter Kazanzides and Nassir Navab. Comprehensive Tracker Based Display Calibration for Holographic Optical See-Through Head-Mounted Display. Submitted to 2017 ISMAR
- [5] Yamauchi, H., Gumhold, S., Zayer, R., & Seidel, H. (2005). Mesh segmentation driven by Gaussian curvature. *The Visual Computer*, 21(8-10), 659-668.
- [6] R. B. Rusu. Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments. PhD thesis, Computer Science department, Technische Universitaet Muenchen, Germany, October 2009.
- [7] Q. Zhan, Y. Liangb, Y. Xiaoa, "Color-based segmentation of point clouds", *Laser Scanning 2009*, vol. IAPRS XXXVIII, Sept. 2009.
- [8] (n.d.). Microsoft Developer. *Spatial mapping*. Retrieved from [http://developer.microsoft.com/en-us/windows/mixed-reality/spatial\\_mapping](http://developer.microsoft.com/en-us/windows/mixed-reality/spatial_mapping)
- [9] C. S. Bamji et al., "A0.13μm CMOS system-on-chip for a 512×424 time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC," *IEEE J. Solid-State Circuits*, vol. 50, no. 1, pp. 303–318, Jan. 2015.
- [10] C. B. Owen, J. Zhou, A. Tang, and F. Xiao. Display-relative calibration for optical see-through head-mounted displays. In Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR), pages 70–78. IEEE, 2004
- [11] M. Tuceryan, Y. Gene, and N. Navab. Single-point active alignment method (SPAAM) for optical see-through HMD calibration for augmented reality. Presence: Teleoperators and Virtual Environments, 11(3):259–276, 2002
- [12] E. Azimi, L. Qian, P. Kazanzides, and N. Navab. Robust optical see-through head-mounted display calibration: Taking anisotropic nature of user interaction errors into account. In Virtual Reality (VR). IEEE, 2017.
- [13] L. Qian, A. Winkler, B. Fuerst, P. Kazanzides, and N. Navab. Modeling physical structure as additional constraints for stereoscopic optical see-through head-mounted display calibration. In IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), pages 154–155. IEEE, 2016.
- [14] L. Qian, A. Winkler, B. Fuerst, P. Kazanzides, and N. Navab. Reduction of interaction space in single point active alignment method for optical see-through head-mounted display calibration. In IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), pages 156–157. IEEE, 2016
- [15] Y. Itoh and G. Klinker. Interaction-free calibration for optical see-through head-mounted displays based on 3d eye localization. In IEEE Symposium on 3D User Interfaces (3DUI), pages 75–82. IEEE, 2014.
- [16] Y. Itoh and G. Klinker. Performance and sensitivity analysis of Indica: Interaction-free display calibration for optical see-through head-

mounted displays. In IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pages 171–176. IEEE, 2014

- [17] *Spectator view*. (2017). *Developer.microsoft.com*. Retrieved 14 May 2017, from [https://developer.microsoft.com/en-us/windows/mixed-reality/spectator\\_view](https://developer.microsoft.com/en-us/windows/mixed-reality/spectator_view)
- [18] Navab, N., Heining, S., & Traub, J. (2010). Camera Augmented Mobile C-Arm (CAMC): Calibration, Accuracy Study, and Clinical Applications. *IEEE Transactions On Medical Imaging*, 29(7), 1412-1423. <http://dx.doi.org/10.1109/tmi.2009.2021947>