Exploring Requirements for Neurosurgical Augmented Reality

Design and Evaluation of an Infrared-Based Inside-Out Tracking Approach on HoloLens 2

Thore Keser[1][0000−0002−9918−692X], Florian Niebling[2][0000−0002−8706−3551], Rahel Schmied-Kowarzik[1][0000−0001−6364−1057], Rebecca Rodeck[1][0009−0008−3572−8890], and Gerko Wende[1]

[1] German Aerospace Center (DLR), Hamburg, Germany
[2] Technische Hochschule Köln, Cologne, Germany

Abstract. Medical professionals require in-situ visualization of X-ray imaging in 3D to assist with surgical navigation and planning. Research on surgical Augmented Reality (AR) focuses on improving the tracking quality of devices and tools but disregards human requirements. Our user study evaluates application areas and potentials for AR in neurosurgery, as well as feedback for HoloLens 2 based inside-out tracking systems. The main findings include a reported simplification in perception and lowered mental workload. Additionally, we propose a simplified implementation for an infrared-based inside-out rigid body pose estimation system on HoloLens 2 to redirect feedback away from the tracking problem and gain access to human requirements for such systems. Decent patient and exemplary instrument tracking are reached.

Keywords: Mixed Reality · Extended Reality · Human-Computer Interaction · Research Mode · Sensors · Contextual Design · User Study.

1 Introduction

Surgical treatment relies on the physician's fine motor skills and the utilized technology. Medical technology is relevant for identifying areas of concern and conducting specific treatments. Identification of tissue and bone structures is done with the help of Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) scans.

In a meeting with staff from the University Clinic Würzburg, neurosurgeons expressed a high mental workload in the interaction with CT images. Images are presented on several flat displays during surgery. Each screen shows the object as a cross-sectional slice image from a different perspective. Additionally, the displays are located aside from the patient, creating an additional rotation offset between the images and the situs. Physicians can use a wheel to scroll through the depth of the images by traversing through individual image slices. This way, medical professionals are forced to put a lot of cognitive effort into understanding the 2D pictures, mapping them back into 3D objects, and mentally aligning them with the patient. Even though CT images are beneficial before and during surgery, they allocate mental resources that could be freed for surgical procedures. Hence, surgeons suggest a digital representation and virtual positioning of CT images to reduce the workload.

Augmented Reality (AR) is a technology that renders digital information with respect to its environment. The technology can create interactive and spatially aligned virtual content. In surgery, AR is able to assist by creating an in-place visualization and thus compensating for relocation and mapping difficulties. Studies show that AR usage can reduce mental load [9] and improve situational awareness in medical (and other) contexts [21].

Aside from tablets and smartphones, a commonly used AR device is Microsoft's HoloLens 2 Mixed Reality Head Mounted Display (HMD). The HMD could apply to hygiene restrictions in surgeries [4, 6], as it is hands-free. HMDs can place virtual content closer to or inside the patient, whereas conventional instruments, tablets, smartphones, or documentation would keep a distance to prevent infections. The overlapping of real and virtual content can be used for guidance and visualization with limited accuracy.

To create convincing visual effects and alignment with the patient, AR needs to understand the patient's position and orientation (pose). To date, research focuses on improving the alignment by enhancing the tracking quality [19, 3, 6]. Challenges lie in the compensation of the low-quality HoloLens sensors and in the reduction of drift and offset, which are introduced through outside-in tracking systems that are combined with the HoloLens inside-out self-localization system [12, 19]. While tracking precision is important for the technical development of future systems, there is little information about human requirements [4], which provide specific information on the most needed features from the users' side of view. An investigation of relevant factors besides tracking quality is necessary.

A user study on AR for neurosurgery requires a functioning tracking system to implement domain-specific features. Moreover, qualitative study results could draw attention to the potentially poor tracking system. Hence, we aim for a simplified tracking approach with decent quality to unlock qualitative feedback on domain-specific aspects and pull the attention away from the tracking quality. A qualitative user study helps exploring neurosurgery's deep insights and needs.

This paper describes both the implementation of a simplified tracking approach for neurosurgical AR (see Fig. 1) and the qualitative user study to gain expert domain knowledge.

2 Related Work

The literature regarding the use of AR in the neurosurgical domain describes most problems within technical areas, such as tracking of tools, its accuracy, precision and refresh rate [5]. To achieve good performance with the used technology, literature quickly dives into a deeper analysis of the inbuilt components of used devices and tracking approaches.

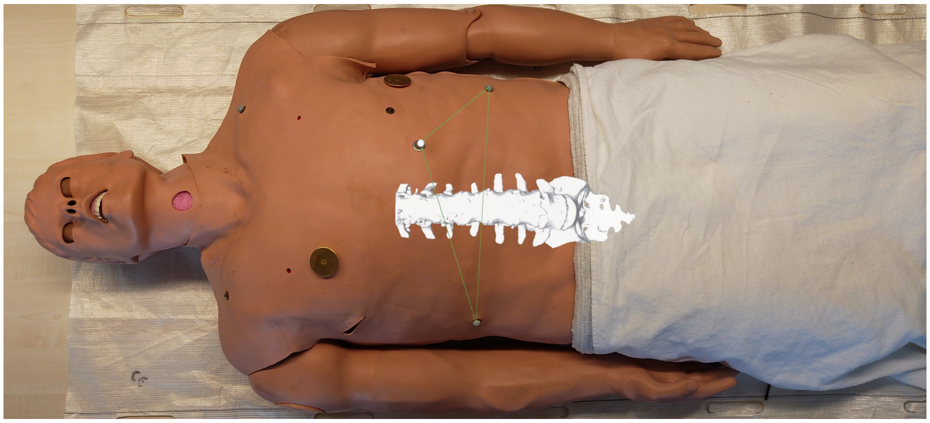

Fig. 1. Infrared target tracking and patient pose estimation for neurosurgical navigation. The green lines visualize the selected anchor points to align the pose of the virtual spine segment with the patient.

Schmalstieg and Hollerer [15] split tracking approaches into two major categories. Outside-in systems use sensors mounted around a tracked object and are strong in detecting position changes whereas the observation of rotation is less precise. Inside-out systems use sensors at the object and observe its surrounding to calculate pose changes of the tracked object. These systems provide improved rotation detection, but are less precise in positioning the object [15].

The HoloLens 2 HMD is an inside-out device, which calculates its own pose using a Simultaneous Localization and Tracking (SLAM) algorithm [20]. When combining the device with external outside-in tracking systems, drift and offset are introduced [17, 19, 12, 3], which is found to originate from the Inertial Measurement Unit (IMU) [19, 12]. Coupling HoloLens to an outside-in system would also limit the developed application to specifically equipped rooms and is thus not preferred. When working in neurosurgical environments, drift and offset need to be minimized. Hence, it is helpful to create a system that tries to overpower device limitations to estimate a more stable pose.

Commonly used tracking approaches rely on fiducial markers, which are fixated at objects to identify them using a camera. For HoloLens, markers with larger diameters are required[3], which might cover relevant working areas in surgery. Furthermore, fiducial markers are rather difficult to detect if the angle to the camera is shallow.

Previous work shows the possibility to access the infrared sensor of the HoloLens 2 HMD. Infrared is already used in current equipment for surgical navigation [13] and is thus unlikely to interfere with medical procedures.

[3] Physical Properties of Image Targets: https://library.vuforia.com/objects/physical-properties-image-based-targets

Infrared approaches promise better tracking performance and require smaller marker spheres.

Kunz et al. [11] utilize the HoloLens Research Mode to access the four visual light cameras and the infrared (IR) sensor. They compare two methods of reflective marker tracking, a) using two of four environmental cameras as a stereo camera setup and b) combining the depth and Active Brightness (AB) stream of the IR sensor. The authors describe relevant steps of the IR tracking implementation. After binarization of the raw image and using a not further specified blob detection algorithm, they apply a HoloLens function, which provides a mapping from frame pixels to world coordinates and vice versa. Position estimation is reached by connecting opposing markers with lines and placing the target center at the cross in the middle. There are no further references on the use of a rigid body pose estimation algorithm or whether a rotation is applied to the target. The authors test the approach by placing the markers on a robotic arm and fixating the HoloLens to a known position. Measuring the actual and virtual movements allows for position comparison between the robot and HoloLens coordinate system. Even though accuracy for the IR setup is slightly lower, they recommend this approach due to the need for an additional light source for the dual camera approach [11].

Another research project [6] also accesses the HoloLens 2 Research Mode to track IR reflective marker spheres. A stereo camera setup with visual light cameras was implemented and the precision optimized by running a custom calibration sequence. When tracking was not available, rigid body movement was predicted using a custom algorithm. The authors tested their approach by comparing the tracked coordinates with a ground truth measurement, which was obtained by an outside-in OptiTrack system. The accuracy correlates with the distance of the target to the device. Without the prediction and thus optimization, they reach refresh rates of 37 Hertz (Hz) [6].

Within most technical approaches, the users' point of view is not evaluated. There is very little information on human interaction that investigates domain insights and user needs. Some research addresses these questions, for example a project [4] that supports surgeons by visualizing the correct screw angle in surgery. Authors highlight the tool as intuitive in use and radiation-free. Limitations are high costs and long setup time, as well as the need to train surgeons before using the headset [4].

A literature review [14] shows that AR can reduce sight deviation and increase the comfort level of team members. Moreover it provides additional information to surgeons, which improves decision making and situational awareness. However, surgeons demand manual activation of functionalities to prevent cluttering and show guidance only if it is required [14]. In general, AR benefits the medical context, especially in surgical workflow improvement and ease of use [1]. Besides clear interfaces and lightweight devices, an unobtrusive patient tracking to increase tolerance against tracking errors is required [1].

The above research describes solutions and psychological factors that are able to benefit the surgical domain. User feedback, which ensures the relevance and priority of solutions and factors for surgeons are missing. Moreover, thresholds are not defined, hence hindering requirement based implementations.

3 Method

We introduce a simplified tracking approach for tool and patient pose estimation as a vehicle to gain access to deeper neurosurgical requirements. As depicted in Table 1, we experimented with different fiducial libraries on HoloLens 2 but rejected many for low performance, such as slow detection or low refresh rate. Some libraries were no longer compilable on current development platforms.

Table 1. Rejected libraries for fiducial tracking using Unity3D.

Library	low performance	not compileable
MRTK QR-Code	×	
Vuforia	×	
ARToolkit	×	
OpenCVForUnity	×	
AprilTag		×
ArUco		×

The introduced approach utilizes the Unity IL2CPP build pipeline to provide a plugin to the Unity3D game engine. The plugin is built into the final HoloLens app. It is derived from a public repository[4] and written in C++ to access relevant sensors through the HoloLens Research Mode API [20]. This project uses the IR sensor in Articulated Hand Tracking (AHAT) mode to gain fast-updating information about the user's close environment. The IR sensor yields two types of raw data: the Active Brightness (AB) image, which is a grayscale image representing the reflectivity of objects, and the depth buffer, which shows the distance of objects to the camera.

Reflective spheres, initially designed for the OptiTrack[5] system, are used to mark real objects. The spheres are brightly visible in the AB images of the IR sensor and are easier to detect in the IR spectrum than with RGB cameras (see Fig. 2). The AB image is probed in a grid pattern to find bright marker blobs. Blob detection is performed on the object reflectivity in the AB image and on depth information to separate blobs that appear behind each other. Blobs are outlined using a contour tracing algorithm [16], which shows to be fast and robust for the occasionally fringed and incomplete images produced by the sensor. The blob centers are retrieved from valid blob contours by calculating the mean coordinates from their contour pixels.

[4] HoloLens2-ResearchMode-Unity on Github: https://github.com/petergu684/HoloLens2-ResearchMode-Unity

[5] OptiTrack Motion Capture Systems: https://www.optitrack.com

The HoloLens Research Mode API provides functions to map image pixels to 3D points from pixel coordinates, depth values, and timestamps. We use this function to transform blob center points to world space coordinates and provide them to Unity in app coordinate space. With Unity, compounds of targets, called multi-targets, are defined and calibrated. Multi-targets represent a group of markers that belong to the same rigid body. The implemented system repeatedly aligns the calibrated multi-target with the reflective marker spheres detected in the surroundings of HoloLens.

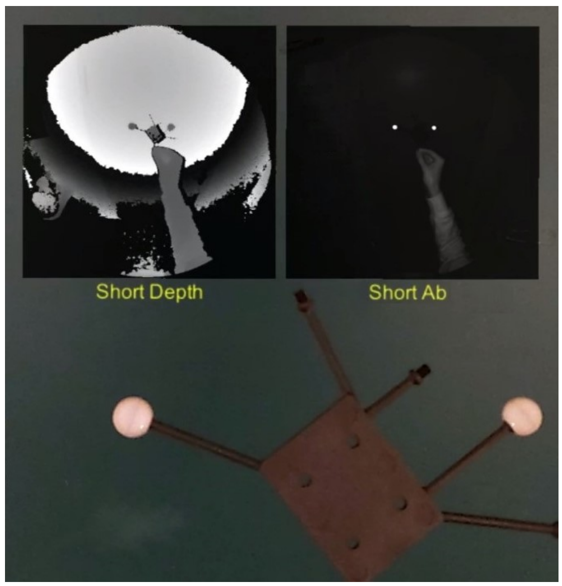

Fig. 2. Depth and AB image from HoloLens 2 sensors.

To align multi-targets with their reflective marker spheres, the system must identify each marker. As markers are equal in size and shape, the only differentiating factor is the set of distances to other markers. To minimize errors, we followed the approach of Gsaxner et al. [6] and determined that every distance within a multi-target must be unique [18]. A marker can be identified by measuring the distances between markers and reconstructing each relative position.

The assignment of a target structure in Unity to a corresponding reflective marker point is calculated using the Hungarian Method [10]. The method is a mathematical approach to solving the assignment problem, which applies to this case. The three best assignments per multi-target are selected by minimum weight [6] and used for the pose calculation of the tracked rigid body.

The pose calculation is performed following Horn's approach [8], where three points define a coordinate space. The pose is expressed as a relative rotational and translational difference, which is applied to the virtual object to align it with its real counterpart.

The introduced approach reduces resource usage for low latency and high performance. Even though the assignment calculation is fast for small point clouds, it still overflows the performance capabilities of HoloLens. Hence, a between-frames assignment, based on close-by pixel movement, is executed to reduce the computational footprint further.

3.1 Study Design

The created tracking system and its application were tested with staff from the University Clinic Würzburg. In a cost-benefit consideration, qualitative interviews are expected to provide the most comprehensive data range with negligible costs. A qualitative evaluation is able to explore undefined contexts. Hence, we chose a semi-structured interview coupled with an activity phase and the Think Aloud [2] method to collect unfiltered and domain-specific data. The collected data was evaluated utilizing methods from the User-Centered or Contextual Design [7] process. The process includes guidelines for qualitative data preparation and consolidation using the Affinity Diagram [7]. The method is qualified for user observations and interviews and is intended to derive design ideas for agile prototyping and software development.

Fig. 3. Participant testing the prototype in the simulator room.

In preparation for the evaluation, key areas of interest were identified. These include the practicability of the system and its ability to integrate into the surgical domain. Moreover, the system's usability is chosen to evaluate the simplicity and value of support for surgeons. Also, measures of quality are intended to answer questions about accuracy, precision, and perceived fidelity. Our interview aims to extract requirements for neurosurgical Augmented Reality.

The interview was scheduled for five working days in a medical simulation room (see Fig 3) and conducted with the knowledge and cooperation of supervisors and responsible medical directors. Interested staff could drop in and participate to prevent disturbance in medical acute care. Individual sessions took about 30 minutes and were carried out in German language.

Each session consisted of three parts. First, several questions about the participants' demographic data, experience, and enthusiasm regarding new technologies were asked. Participants were briefly introduced to the HoloLens device to ensure proper fit during testing. With an introduction and reminder about the Think Aloud procedure, they were exposed to the demo application. During the second phase, participants were given small tasks to try out the application and improve the user's awareness and perception of all virtual elements in the scene. Tasks included implicit requirements for position and view direction changes, patient movement, and clipping tool interactions (see Fig. 4). Users were instructed to take their time to test all imaginable prototype features before requesting new tasks incentives from the study supervisor. After completing all tasks, participants stopped the testing and proceeded to the interview (third phase).

The interview was recorded and transcribed. Written sentences were shortened, and information that didn't contain relevant aspects to the study was excluded. Participant data was anonymized with the change from audio to text format. We extracted and interpreted key information from the transcribed interview answers according to the Contextual Design method by Holtzblatt, Wendell, and Wood [7]. The captured data was written onto Affinity notes and consolidated into an Affinity Diagram to provide insights into the user's context [7]. Key information was rearranged to represent the main factors of experience, quality, practicability, and usability.

4 Results

The created application incorporated a sample spine segment aligned with the patient's pose using the introduced tracking system (see Fig. 1). The system is able to track the patient's pose and a hand-held clipping tool (see Fig. 4) while maintaining a low computational footprint. Compared to previous work, which reached five and 55 to 60 frames per second (FPS), respectively [12, 11], the system updates the pose with 45 FPS, originating from the infrared sensor refresh rate. We additionally achieved Unity3D integration with an app performance of 59 to 60 FPS on the HoloLens 2 HMD. The tracking quality is decent and drift-free through direct coupling with the HoloLens built-in environment tracking.

The system acts as an enabler for staff imagination and requirement generation. The semi-structured interview provides rich feedback and shows insights into challenges and procedures within the neurosurgical domain.

A total of five participants (U1 - U5) attended the interviews. Four identified as male and one as female. The age of the participants ranged from 25 to 37 (M=31.2 SD=4.0). Two subjects were students without chosen specialization. Three were specialized orthopedics and/or trauma surgeons with experience in neurosurgery. The students reported being in 8th and 13th semesters, while the

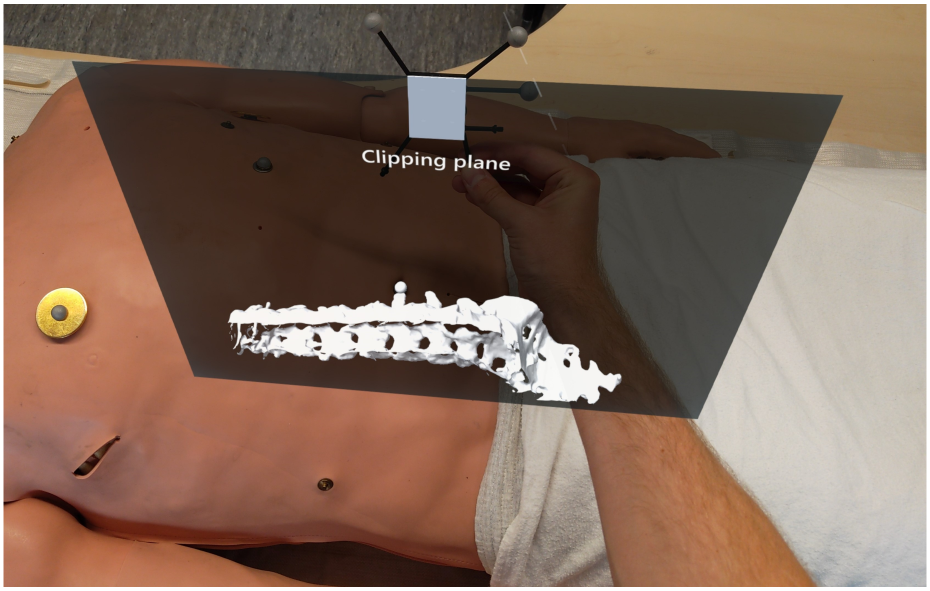

Fig. 4. Sample application using the implemented tracking approach to track a mannequin patient and sample tool to clip the virtual spine segment.

professionals reported a work experience between four and nine years (M=5.7 SD=2.4). One of five participants reported a light debility of sight, which wasn't corrected by wearing glasses or contact lenses while using the AR HMD. Also, one participant was affected by red-green blindness.

4.1 Qualitative Feedback

In this section, we summarize the gathered user feedback. The qualitative data has been rearranged and interpreted to compress the information and identify relevant factors. In total, 23 data groups were created, which fit into five major categories. Please note that user quotes have been translated into English and do not contain literal speech.

Augmented Reality in neurosurgery. Interview participants see great advantages to the use of AR technology in neurosurgery. *It reduces distractions by blending out irrelevant objects and other people. It emphasizes on the shown content* (U3).

Perspective-aligned renderings are currently uncommon for the domain. *The spine is visualized in perspective* (U2) expresses the perception of patient-aligned content that follows real-world perspective cues. *I think this helps with understanding the situation* (U1).

Users are also amazed about the interaction methods that have become possible with the technology. When reporting ideas of separating holographic content and zooming into detailed areas, medical professionals often showed gestures with their hands to explain the intended interaction technique. *I could divide the hologram and look at each part separately* (U2). When separating the hologram, users could take it out of the aligned pose to drag it closer or above the patient. Users expect that dragging the hologram back into the patient will snap it back into place, where it is correctly aligned and reacts to tracking changes.

It is an obligation to keep the surgical area (situs) clear of objects to prevent infections. However, *with HoloLens I could place X-ray images onto the patient* (U1) because virtual images are sterile by default. With an open situs, the surgeon's hands are used to feel the way for navigation and planning. However, *for the interaction with HoloLens, I do not need to touch anything* (U1). Thus, the device would remain uncontaminated.

It could have been more apparent to some users that the tracking system solely relies on HoloLens' inbuilt sensors. Line-of-sight errors were anticipated as current navigation systems are based on fixture-mounted sensors. *Where is the sensor?* (U3) was asked to prevent tracking issues. The independence of other devices is positively mentioned.

Some feedback attributed to the HoloLens 2 device itself is collected. *A larger visor and visualization area would be beneficial* (U3). The device's weight is reported to be acceptable, and the overall comfort is good. For tall people, more than the one-meter AHAT tracking distance is required. *I need to take the head down. In this position, I would not operate on patients.* (U3). In general, the requirements for the maximum tracking distance of the system are expressed to be within one arm's distance. *Usually, I stand with my thighs against the operating table when conducting surgery* (U1). Larger distances can be useful in multi-user environments where assistants or students observe.

Tracking approaches. Feedback was also given concerning the currently used surgical navigation system. It is able to combine information from medical imaging devices with the position of tracked instruments and patient. The system consists of a sensor arm, which tracks infrared markers. Results are visualized on screens aside from the patient. As the system provides only 2D visualizations, different viewing angles are presented aside from each other in a split screen layout. Users described the system as helpful but rather challenging to handle. *The positioning of the camera is error-prone* (U3) and *the system faults when markers are moved.* (U1). *It takes a lot of time until the system is set up and ready to use* (U1).

The data shows a difference between tracking of the clipping tool and the patient for the implemented tracking approach, even though the implementation for both features is the same. Differences originate from different smoothing settings that allow a faster but less smooth position change for the clipping tool but slower and smoother spine alignment.

For patient tracking, users noticed that they had to stay in a specific range. *From further away, the tracking was less sophisticated* (U4). Also, small movements did not take effect. The latter appears if movements are tiny so that adjustments are not seen or not even recognized by the system. Moreover, strong smoothing created a perceived delay in position updates. *The spine could move faster, so I don't have to wait for it to adjust* (U4). *For surgery, the delay is too long. When I move the patient, the visualization has to act faster.* (U1).

The tracked area of the clipping tool was not limited (U2). As the tool stays in the surgeon's hands, the range of the AHAT sensor is sufficient for this task. However, *The tool must be oriented with a specific angle to the camera* (U3). This issue could be fixed by adding another marker to the tool. *For placing corrective screws during surgery, the tool is not precise enough* (U5). Regarding delay, the clipping tool is perceived more positively than patient tracking. Most users rate the current delay as not noticeable, reasonable, or adequate for the sample task. *There is too much delay for surgery, but it would be okay for diagnostics* (U2).

Requirements proposed by participants can be grouped into translational and rotational precision. More precise tools are expected to improve patient outcomes in surgery. *In surgery, every millimeter is crucial* (U1). Demands also vary based on experience level. More experienced surgeons require higher precision, while students estimate lower values. Moreover, *the tolerance for deviation during surgery is higher at the lower spine and reduces with the size of the vertebral bodies* (U3). The requirements for translational precision and accuracy are stated by multiple participants to be below one millimeter and one degree for rotational deviations.

System quality. The overall visualization quality and level of detail of the spine model are delicate enough for medical professionals to name various areas of the spine. However, surgeons state that they *would need a more detailed hologram to work with it in surgery* (U2). Some users rate the image resolution sufficient, while others demand higher resolutions for detailed inspection. Regarding hologram size and proportions, surgeons trust the application to show a correct scale.

The colors of the selected spine structure are almost well selected. Even if nobody reported difficulties with the selected color, medical professionals know the shades of different structures from educational books and personal experience. They stated to prefer applications to respect these color schemes.

The HMD's brightness is disturbing when trying to see details in the hologram from shades. *I believe that the visualization could dazzle in surgery* (U1). As HoloLens has a physical button for brightness settings, this would not be an issue as long as gestures are available to manipulate the setting during surgery. However, contrast is lost when the HMD adds less light to the scene, and a more substantial transparency effect might occur.

All experiences were positively denoted. Differences are in affective emotions that show surprise and fascination, as well as retrospective reflections summarizing the experience as helpful. Many positive emotions can be attributed to the first contact with AR, which makes the experience exciting and new.

Additional feedback can be grouped into Usability measures defined by DIN EN ISO 9241-11, which consist of effectiveness, efficiency, and satisfaction. For effectiveness, users stated that the tool helps them reach their goals, e.g., *I was able to take a look into the spinal canal* (U5). With efficiency, users report that interactions happen fast, and the scene changes according to their position in the room. *The clipping of the spine happens in real-time. I really like it* (U2). Users were also satisfied with the provided interaction techniques, audio and visual feedback, and User Interface (UI) elements.

With the introduced application, visualizations are in 3D, which was attributed to being helpful for the understanding process. *With X-Ray images, the visualization is rather confined* (U2) and *X-Ray images are more precise, but this system is simpler because it is in three dimensions* (U5). Users reported the application to be *very easy* (U1) to use with no requirement for additional guidelines. All visualizations appear real and correct, as one *would expect from the real world* (U4). This corresponds to the user's reality perception and lowers the mental load.

Application areas. This section summarizes application areas for the system and technology within the medical domain. A central area of work is surgery, in which medical professionals wish more support for navigation and confirmation of correct placement. *By showing an angle or an axis, the system could improve screw and plate placement in surgery* (U2). With the hypothetical possibility of connecting HoloLens with a patient data platform, it is possible to quickly view and exchange data without requiring an office space. *With access to my desk, I could view relevant reports, documents, and X-ray images during surgery and in the ward* (U1). Moreover, users anticipate the technology to be beneficial in orthopedics, vascular surgery, or other medical areas, such as psychology.

For complex disease patterns, collaboration is required. As virtual content on HoloLens is only visible to one user, professional communication and exchange are hindered. Hence, the need to share the view and visual content and the ability to interact with a virtual scene in a multi-user scenario is required. *After primary treatment, we could discuss a complex trauma in our team* (U2). With the help of AR, viewing content is not restricted to a single device, but every user could see their image and communicate with others. This would also fix the rotational offset problem, which has been a central goal for this work.

Students who attend the surgery for educational purposes examine the situs and all relevant information. *Currently, all students must gather around one display* (U1), which limits class sizes and requires students to switch places for better views. Augmented Reality in educational contexts increases the ability to understand and reconstruct three-dimensional shapes. Students are equipped with many images but need to combine them to understand the visualized content. *Using a book, I will need at least five images to understand a situation. With this technology, it would be much simpler* (U5).

In anatomy classes, the utilization of donated bodies is common. These are limited resources, as every cut irreversibly changes the sample. Also, preserving the bodies is costly. *With this system, we could repeat certain scenarios* (U2) and reduce resource consumption. With the ability to virtually create content and artificially track movements, it is possible to create training applications that allow the simulation of surgical procedures without costly resources. *We could simulate surgeries and train our cutting techniques* (U1).

System potential. Participants stated that *the technology has high potential* (U1) within the medical domain. *The technology can improve and speed up our processes* (U4). *With the currently existing staff shortage, technology and digitalization can help to utilize the resources with the existing staff more efficiently* (U1). This would also result in a less stressful workplace and better care.

The patient outcome, which describes the treatment quality, speed, and success of therapy, remains the primary goal for medical professionals. This is the main reason for rejecting technological advances if they are more complex than helpful or do not speed up the process. *New technology must be worthwhile to become helpful* (U2). Current systems are trustworthy and create an appreciated assistance within the surgical domain. Medical devices are trusted to create little to no errors by maintaining high standards.

When referring to the implemented system and application, users are willing to use it in real scenarios. *If patient and setting allow it, one could try the system with a rather simple diagnosis. However, I would not trust it yet* (U3). The reason for the experimental initiative is the hope that the technology and application will solve long-term goals with further development and research. *In the long run, I believe this project will be very beneficial for us* (U3).

At the same time, there are great concerns about the system's usefulness. *I do not believe that we can resign from the current imaging system* (U4). *With the current situation, I would rely on existing systems* (U1). As the current tools are well proven, a transition period to the new system would require parallel coverage with traditional systems for a long time.

Radiation imagery such as CT or X-ray is a helpful and often used tool in preparation and surgery. As radiation potentially harms human bodies, its use is carefully balanced between patient benefits and disadvantages. *Daily, we face the risks of radiation* (U3), as it also affects staff in the operating room. Every additional image adds to that risk, and consecutive images must be taken to confirm the current position or placement of tools. If systems align virtual content with the patient, react to position changes, and are interactable, they increase the potential to reduce radiation in surgery.

5 Discussion

An infrared-based tracking approach for neurosurgical AR was achieved and comprises several factors. These include patient tracking, which is required for augmentation in object (patient) coordinate space, instrument tracking for guidance visualizations, infrared light technology for more straightforward marker detection, and efficient marker assignment. These components provide a simplified solution and implementation for the complex idea of neurosurgical task-related AR.

The IR sensor is the fastest but not the most accurate sensor. Marker detection is simpler to implement, as binarization quickly yields the marker positions in camera image coordinates. With the depth stream of the IR sensor and HoloLens internal functions, these coordinates can be joined with depth information and projected into 3D space. As HoloLens is aware of its environment, the calculated world coordinates align with the refined SLAM-based world and app coordinate systems. Our solution is drift-free and performant on HoloLens.

The qualitative evaluation is able to provide rich feedback and shows excellent insights into challenges and work within the neurosurgical domain. It remains arguable if more participants would show more aspects of the surgical domain regarding the implemented system. More extended opportunities to attend the interview sessions might not yield more participants due to their high workload and sudden medical events. More participants are not expected to enlarge the picture significantly but rather strengthen the results by emphasizing individual factors.

Within the evaluation, novice and experienced medical professionals showed different results, because novices might not have experience in surgery or knowledge about specific procedures and tools. Thus, the given feedback sometimes shows the gap in experience through the ability to answer specific questions, as well as differing impressions about the system and stated requirements.

Using the qualitative approach, we extracted requirements for neurosurgical Augmented Reality, which applications should fulfill to be of value for the domain. Applications must:

- fulfill usability measures in effectivity, efficiency, and satisfaction.
- be self-descriptive with its UI elements and interactions.
- solve a specific problem that is relevant to the domain.
- respect expectations and experiences from the surgeon's real world.
- provide a transition period and fallback technology on system introduction.
- minimize the time required to set up and launch the system.
- render virtual content in perspective with respect to the patient's pose.
- remain sterile in interaction and device usage.
- be able to place virtual content close or into the situs.
- be self-contained and independent from external sensors.
- provide trackable space in arm-length distance.
- be able to track objects outside of the optimal range (with less precision).
- reach 0.5 to 1 mm translational precision.
- reach below 1° rotational precision.
- reach firm real-time updates of the tracked position (at least 45 FPS).
- provide a resolution that can show X-ray grade details.
- show detailed 3D visualizations.

6 Conclusion

We introduced a simplified tracking approach on HoloLens 2 to unlock domain-specific insights from qualitative interviews. An infrared-based inside-out rigid body pose estimation system has been implemented. The system utilizes the HoloLens Research Mode to access the IR sensor and provided API methods for reflective marker sphere tracking and coordinate system transformation. By running a fast contour tracing algorithm, the marker sphere centers are extracted and assigned to tracking targets using the Hungarian Method. A pose for the tracked patient and tool is calculated from the three best assignments.

The system addresses problems that originate form technical limitations. For example, in surgery space around the patient is occupied and hinders conventional approaches in establishing a line of sight to the tracked object. Using the inside-out approach, the system is able to provide a drift-free augmentation in close range using markers that do not impede the small augmentation area. As processing power on the HoloLens 2 device is limited, the system remains a low computational footprint reach an experience with reduced latency.

Surgeons reported lower mental workload, which was contributed by the 3D reconstruction of the medical image and simultaneous patient-aligned positioning of the reconstruction. Medical professionals focus on central tasks and discard distractions, which they attribute to the illumination effect of the visualization. Moreover, hands-free interaction and sterility are rated to be beneficial.

The feedback concerning the future use of the system was very positive. With the knowledge of close-by critical tissues or blood vessels, delicate tool positioning and path planning are possible. Visual guidance in tool positioning in relation to the situs would also reduce the need for radiation during surgery. Additionally, multi-user visualization can improve training, where the space around the patient is limited. At the same time, team collaboration during surgery can be improved.

Our system acts as an enabler for staff imagination and requirement generation. The Affinity Diagram provides rich feedback and opens great insights into challenges and procedures within the neurosurgical domain. Medical professionals hope the system revolutionizes their work in simplicity, flexibility, surgical planning, and navigation.

References

1. Burström, G., Persson, O., Edström, E., Elmi-Terander, A.: Augmented reality navigation in spine surgery: a systematic review. Acta Neurochirurgica **163**, 843–852 (2021). https://doi.org/10.1007/s00701-021-04708-3
2. Charters, E.: The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Education Journal **12**(2) (2003). https://doi.org/10.26522/brocked.v12i2.38
3. Doughty, M., Ghugre, N.R., Wright, G.A.: Augmenting performance: A systematic review of optical see-through head-mounted displays in surgery. Journal of Imaging **8**(7), 203 (2022)
4. Farshad, M., Fürnstahl, P., Spirig, J.M.: First in man in-situ augmented reality pedicle screw navigation. North American Spine Society Journal (NASSJ) **6**, 100065 (2021). https://doi.org/10.1016/j.xnsj.2021.100065

5. Fida, B., Cutolo, F., di Franco, G., Ferrari, M., Ferrari, V.: Augmented reality in open surgery. Updates in surgery **70**(3), 389–400 (2018). https://doi.org/10.1007/s13304-018-0567-8
6. Gsaxner, C., Li, J., Pepe, A., Schmalstieg, D., Egger, J.: Inside-out instrument tracking for surgical navigation in augmented reality. In: Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology. pp. 1–11 (2021). https://doi.org/10.1145/3489849.3489863
7. Holtzblatt, K., Wendell, J.B., Wood, S.: Rapid contextual design: a how-to guide to key techniques for user-centered design. Elsevier (2004)
8. Horn, B.K.: Closed-form solution of absolute orientation using unit quaternions. Josa a **4**(4), 629–642 (1987). https://doi.org/10.1364/JOSAA.4.000629
9. Jeffri, N.F.S., Rambli, D.R.A.: A review of augmented reality systems and their effects on mental workload and task performance. Heliyon **7**(3) (2021). https://doi.org/10.1016/j.heliyon.2021.e06277
10. Kuhn, H.W.: The hungarian method for the assignment problem. Naval research logistics quarterly **2**(1-2), 83–97 (1955). https://doi.org/10.1002/nav.3800020109
11. Kunz, C., Maurer, P., Kees, F., Henrich, P., Marzi, C., Hlaváč, M., Schneider, M., Mathis-Ullrich, F.: Infrared marker tracking with the hololens for neurosurgical interventions. Current Directions in Biomedical Engineering **6**(1) (2020). https://doi.org/10.1515/cdbme-2020-0027
12. Lee, S., Jung, H., Lee, E., Jung, Y., Kim, S.T.: A preliminary work: Mixed reality-integrated computer-aided surgical navigation system for paranasal sinus surgery using microsoft hololens 2. In: Computer Graphics International Conference. pp. 633–641. Springer (2021). https://doi.org/10.1007/978-3-030-89029-2_47
13. Mezger, U., Jendrewski, C., Bartels, M.: Navigation in surgery. Langenbeck's archives of surgery **398**(4), 501–514 (2013). https://doi.org/10.1007/s00423-013-1059-4
14. Qian, L., Wu, J.Y., DiMaio, S.P., Navab, N., Kazanzides, P.: A review of augmented reality in robotic-assisted surgery. IEEE Transactions on Medical Robotics and Bionics **2**(1), 1–16 (2019). https://doi.org/10.1109/TMRB.2019.2957061
15. Schmalstieg, D., Hollerer, T.: Augmented reality: principles and practice. Addison-Wesley Professional (2016)
16. Seo, J., Chae, S., Shim, J., Kim, D., Cheong, C., Han, T.D.: Fast contour-tracing algorithm based on a pixel-following method for image sensors. Sensors **16**(3), 353 (2016). https://doi.org/10.3390/s16030353
17. Sitole, S.P., LaPre, A.K., Sup, F.C.: Application and evaluation of lighthouse technology for precision motion capture. IEEE Sensors Journal **20**(15), 8576–8585 (2020). https://doi.org/10.1109/JSEN.2020.2983933
18. Steinicke, F., Jansen, C.P., Hinrichs, K.H., Vahrenhold, J., Schwald, B.: Generating optimized marker-based rigid bodies for optical tracking systems. In: VISAPP (2). pp. 387–395 (2007)
19. Tu, P., Gao, Y., Lungu, A.J., Li, D., Wang, H., Chen, X.: Augmented reality based navigation for distal interlocking of intramedullary nails utilizing microsoft hololens 2. Computers in Biology and Medicine **133**, 104402 (2021). https://doi.org/10.1016/j.compbiomed.2021.104402
20. Ungureanu, D., Bogo, F., Galliani, S., Sama, P., Duan, X., Meekhof, C., Stühmer, J., Cashman, T.J., Tekin, B., Schönberger, J.L., et al.: Hololens 2 research mode as a tool for computer vision research. arXiv preprint arXiv:2008.11239 (2020). https://doi.org/10.48550/arXiv.2008.11239

21. Woodward, J., Ruiz, J.: Analytic review of using augmented reality for situational awareness. IEEE Transactions on Visualization and Computer Graphics **29**(4), 2166–2183 (2022). https://doi.org/10.1109/TVCG.2022.3141585