Comparison of Augmented Reality Display Techniques to Support Medical Needle Insertion

Florian Heinrich, Luisa Schwenderling, Fabian Joeres, Kai Lawonn and Christian Hansen

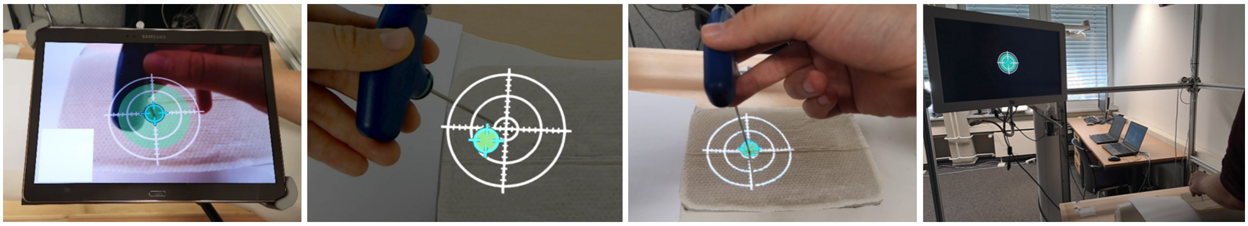

Fig. 1. Needle navigation aid visualized using different display techniques. a) Video see-through visualization on a stationary tablet computer. b) Visualization displayed by an optical see-through head-mounted display. c) Spatial augmented reality visualization directly projected onto a phantom. d) Visualization shown on a monitor.

Abstract—Augmented reality (AR) may be a useful technique to overcome issues of conventionally used navigation systems supporting medical needle insertions, like increased mental workload and complicated hand-eye coordination. Previous research primarily focused on the development of AR navigation systems designed for specific displaying devices, but differences between employed methods have not been investigated before. To this end, a user study involving a needle insertion task was conducted comparing different AR display techniques with a monitor-based approach as baseline condition for the visualization of navigation information. A video see-through stationary display, an optical see-through head-mounted display and a spatial AR projector-camera-system were investigated in this comparison. Results suggest advantages of using projected navigation information in terms of lower task completion time, lower angular deviation and affirmative subjective participant feedback. Techniques requiring the intermediate view on screens, i.e. the stationary display and the baseline condition, showed less favorable results. Thus, benefits of providing AR navigation information compared to a conventionally used method could be identified. Significant objective measures results, as well as an identification of advantages and disadvantages of individual display techniques contribute to the development and design of improved needle navigation systems.

Index Terms—Medical augmented reality, display techniques, surgical navigation systems, needle guidance, visuospatial task

✦

1 INTRODUCTION

Compared to open surgery, minimally invasive interventions do not require the patient to be opened, which results in benefits like reduced risk of infections and less damage to healthy tissue [7, 30]. Instead, image guidance methods like radiological imaging are employed to compensate for missing visual and haptic feedback [43]. Treatments like tumor ablations or tissue biopsies require the precise insertion of needle-shaped instruments. Surgical navigation systems have been developed to further support performing interventionists [40]. By providing visualizations guiding the needle insertion process, these systems were shown to reduce the procedure time and the number of needed imaging scans, as well as to improve targeting accuracy [32, 36]. Guidance information of conventional navigation systems is usually displayed on external monitors. Frequently consulting these screens interrupts the interventionalist's attention to the patient and increases time pressure and mental workload [33]. Additionally, the spatial separation of needed visual information and the intervention site results in a complicated hand-eye-coordination [5, 11]. In previous work, the concept of augmented reality (AR) was applied to surgical navigation systems to mitigate these issues by directly superimposing the view on the patient with needed information [14].

Such systems were developed based on different display techniques. In the literature, these techniques are either classified by their augmentation type, i.e. video see-through (VST), optical see-through (OST) and spatial AR [39], or their display device type, i.e. head-mounted displays (HMD), handheld devices and projector-camera-systems [28, 38]. The terms spatial AR and projector-based AR are often used synonymously. However, spatial AR is also used to describe VST or OST stationary displays [8]. AR navigation systems have been developed fitting into each of these categories. Das et al. [10] and Wacker et al. [46] described VST HMD navigation systems. Suitable visualizations were superimposed on camera views and displayed by head-worn stereoscopic displays. Both works were evaluated by measuring the accuracy of guided needle insertions in radiological images. Bork et al. [4] and Seitel et al. [44] developed VST navigation systems displaying a camera view on the intervention site on a stationary monitor. Images were augmented with visualizations guiding the insertion process. Needle insertion accuracies were measured to assess both systems' efficacy. Hecht et al. [18] applied VST techniques to a handheld smartphone device by superimposing the internal camera view with needle alignment information. They compared their approach with a standard CT-guided method (i.e. without navigation system) and reported favorable results for the AR solution. Most of related work investigating OST AR navigation systems used HMDs to visualize guidance information within the view of the user [1, 15, 20]. These systems were positively assessed by measuring accuracy data of guided needle insertion tasks. Fritz et al. [13] described a spatial OST monitor positioned in front of an MRI bore to navigate needle insertions. Absolute needle placement errors were evaluated in a phantom study. Efforts regarding projector-camera-systems were mostly made using spatial AR systems directly augmenting the patient with navigation information. Krempien et al. [27] evaluated

---

- *F. Heinrich, L. Schwenderling, F. Joeres and C. Hansen are with the University of Magdeburg, Germany, and Research Campus STIMULATE. E-mail: hansen@isg.cs.uni-magdeburg.de.*
- *K. Lawonn is with the University of Jena . E-mail: kai.lawonn@uni-jena.de.*

Manuscript received 18 May 2020; revised 26 July 2020; accepted 17 Aug. 2020.
Date of publication 2 Oct. 2020; date of current version 3 Nov. 2020.
Digital Object Identifier no. 10.1109/TVCG.2020.3023637

such an approach in a clinical study measuring absolute instrument insertion accuracy. Mewes et al. [35] described the integration of a spatial projector-camera-system into an MRI bore and Heinrich et al. [19] evaluated different navigation visualizations for projected needle navigation. Gavaghan et al. [14] developed a handheld projector to spatially augment the patient with navigation aids and assessed their approach qualitatively.

While these various navigation systems employing different types of AR display techniques were evaluated individually for visualizing instrument guidance information, a comparison between displaying systems was not conducted before in this domain. However, because of different displaying properties, user performance and perception may differ depending on used devices. In the literature, VST and OST HMDs have been compared for different use cases. Rolland et al. [42] investigated these systems for medical visualization. They concluded, that OST devices experienced less real world input latency and occlusion, leading to better patient safety. However, VST devices enabled an improved view on the virtual content. Juan et al. [26] evaluated VST and OST HMDs for animal phobia treatment and argued, that VST devices induced a greater sense of presence. Ballestin et al. [2] investigated both display techniques' effects on the perception of the user's peripersonal space. In their experiments, the OST HMD enabled a better depth perception. A user study evaluating a broader set of device types was conducted by Hald et al. [17]. They compared a VST stationary display (i.e. a mounted tablet computer), an OST HMD, a projector-camera-system and a graphical overlay on a monitor for the visualization of sub-surface positions. In the experiment, a pointer tip had to be moved to these positions. The projection-based system achieved higher user satisfaction scores and required less task completion time. However, these results are not specific to applications in the medical domain.

This work aims at investigating similar effects for the visualization of AR needle navigation aids. To this end, a conventionally used monitor-based method is compared with a VST stationary display, an OST HMD and an spatial AR projector-camera-system. All systems were used to display the same navigation visualization. Effects of different registration accuracies were neglected measuring insertion accuracy using virtual tracking data instead of absolute positions of the real instrument. Moreover, no absolute targeting accuracy was measured. Instead, individual angular and depth deviation parameters were regarded to mitigate effects of unrealistically stiff needles and plastic phantom deformations. The resulting prototypes are presented in the supplemented video.

2 MATERIAL AND METHODS

A within-subject design user study was conducted to investigate effects of different display techniques on user performance in navigated needle insertion tasks. In the following, details and rationales of the experiment are given.

2.1 Apparatus

All prototypes have been developed using the game engine Unity (Unity Technologies, USA). Needle positioning data was calculated by a separate base application using optical tracking data (fusionTrack 500, Atracsys LLC, Switzerland). This data was registered to a common world coordinate system using fiducial markers at a known position. During the study, needles were inserted into a human torso phantom filled with candle gel covered by a paper towel. A digital surface scan in the world coordinate system was acquired using a photogrammetric measurement system (ProjectionTools, domeprojections.com GmbH, Germany). Subsequently rendered navigation visualizations were based on intersection calculations between this surface and the instrument tracking data. Needed information was transmitted wirelessly to the respective output devices.

To guide the needle insertion process, a crosshairs-shaped visualization was adapted. The concept was evaluated best for projector-based AR by Heinrich et al. [19] and is also used by commercially available monitor-based navigation systems [40]. Fig. 2 illustrates the navigation aid. The center of the crosshairs marks the planned insertion site. A small circular glyph, representing the needle handle, encodes needle orientation information. The glyph needs to be aligned at the crosshairs center for the instrument to be oriented correctly. By inserting the needle, a circular filling increases in diameter, thus representing insertion depth. When the filling completes the crosshairs grid, the needle tip is as close to the planned target position as possible based on the current instrument orientation. Angle and depth accuracies were additionally mapped to the colors of the respective indicators. In contrast to Heinrich et al. [19], these colors were based on a single hue color scale with varying luminance levels instead of using a traffic light metaphor with multiple hues, because quantitative ordered data, e.g. accuracy values, is better encoded by saturation or luminance color channels than by the hue channel [37]. Discrete colors were assigned to four accuracy levels defined by the thresholds 15 mm / 4.5°, 3.33 mm / 1° and 1 mm / 0.3°.

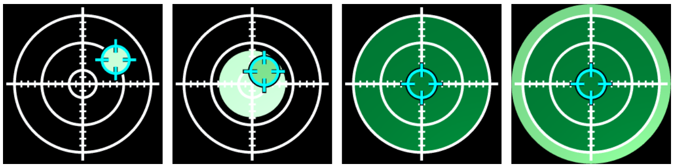

Fig. 2. Navigation visualization. A colored circle (blue border) represents needle orientation accuracy. Insertion depth is encoded by the radius of the circular filling. From left to right the orientation gets improved and the needle is inserted further. The last image shows a needle being inserted too far.

2.2 Display Techniques

For this work, three AR display techniques were selected based on combinations of classifications by augmentation type [39] and display device type [38], that are most commonly used in related work. Handheld devices were not included in this experiment, because they are less comparable to the other device types due to one hand being occupied. This may restrict movement ergonomics and complicate switching the instrument between hands.

2.2.1 Video See-Through AR: Stationary Display

In related work, different device types were used for VST AR navigation systems. For this experiment, a stationary display approach was selected, because VST HMDs were already evaluated as less well suited in the medical domain compared to OST HMDs [42] and handheld devices were excluded for the study. Similar to the work of Marien et al. [34], a tablet computer was mounted obliquely above the phantom with its front facing camera capturing the insertion site (see Fig. 3). Based on wirelessly received needle positioning information, the navigation visualization was rendered perspectively correctly at the respective planned insertion positions, which is shown in Fig. 1a). This was realized by an initial registration step using the Vuforia AR SDK (PTC Inc., USA) to track an image marker placed on the phantom. The marker was attached to a reference frame including fiducial markers tracked by the optical tracking camera, thus enabling the calculation of needed coordinate transformations. Because of a fixed mounting position, this step needed to be performed only once at program start. A transparency effect was applied to the virtual rendering before superimposing the camera view, to facilitate the perception of the real instrument.

2.2.2 Optical See-Through AR: Head-Mounted Display

Most of the past approaches on developing OST AR navigation systems concentrated on using HMDs. Therefore, this display technique was also included in the experiment. Qian et al. [41] compared different OST HMDs for surgical interventions and concluded, that the mixed reality glasses Microsoft HoloLens (first generation) were best suited among evaluated devices. Hence, this work's OST prototype was also developed for this device. Wirelessly received instrument tracking data was transformed into the HoloLens' coordinate system using a similar approach as for the tablet computer. The Vuforia AR SDK

Fig. 3. Study apparatus. VST tablet computer, OST HMD, projector-camera-system and monitor were used to display navigation aids. An optical tracking camera was used to obtain needle position data and for phantom registration. Only the projector opposite from the participant was used (top left image corner).

detected the image marker on the reference frame including fiducial optical tracking markers and calculated its transformation matrix in the HoloLens' space. Because of the known relations between both markers on the reference frame, a transformation between HoloLens and world coordinate system could then be calculated. The registration step needed to be performed only once at program start because of the AR glasses' self-localization algorithms [45]. Respective navigation visualizations could then be rendered at the correct positions on the phantom. The HoloLens' brightness value was adjusted to amplify the device-specific transparency effect of rendered contents, so that the held needle was not occluded. Fig. 1b) shows a view through the OST HMD captured by Microsoft's Mixed Reality Capture software. The resulting image's transparency and brightness properties differ from the view actually perceived by the users.

2.2.3 Spatial AR: Projector-Camera-System

Besides VST and OST AR display techniques, spatial AR projector-camera-systems have been used, to visualize instrument navigation aids directly on the patient. Therefore, this type of technique was also investigated in this work's comparison. A projector (Barco F22 WUXGA, Barco GmbH, Germany) was mounted obliquely above the phantom's position (see left projector in Fig. 3). Its position and orientation was adjusted to minimize effects of self-shadowing by a user standing on the opposite side of the phantom. The projection system was calibrated with the same software used to obtain the phantom surface scan, resulting in intrinsic and extrinsic projector parameters in the world coordinate system. Using a projection mapping approach, navigation aids could then be projected directly onto the phantom, as visualized in Fig. 1 c).

2.2.4 Non-AR: Monitor-Based Navigation System

Finally, a conventional monitor-based navigation approach was implemented to investigate differences between included AR prototypes and a currently used method. In this condition, the navigation visualization was always shown at the center of a nearby positioned monitor (see Fig. 1d and Fig. 3). The used display was part of a commercially available surgical navigation system (CAS-ONE IR, Cascination AG, Switzerland). Thus, an authentic environment could be simulated. Because the planned insertion site could not be directly visualized by the crosshairs' center position using this method, the navigation aid was extended by a small cross indicating the extended needle tip position on the phantom's surface. That cross needed to be aligned at the crosshairs' center to find the correct insertion position.

2.3 Tasks

The experiments primary task was the insertion of a tracked needle instrument into a phantom. For this, visualized insertion position, angle and depth needed to be adhered as accurately as possible. These parameters were randomly generated before each trial. Insertion depth could range from 70 mm to 90 mm and insertion angles were selected between 65° and 90° at a randomized direction around the insertion position.

For each insertion, participants also needed to complete a visuospatial secondary task to assess mental workload. It was hypothesized, that participants would perform worse in this task, if needles were inserted under more demanding conditions, i.e. that display techniques would require the binding of less mental resources, if subjects were able to solve the secondary task more successful. To this end, a mental rotation task as described by Cooper and Shephard [9] was adopted. Participants were presented a stimulus in form of a tilted alphanumeric character, that needed to be mentally rotated to an upright position to decide whether the character was displayed normally or mirror-reversed. When identifying normal characters, a button on a hand-held remote control needed to be pressed as quickly as possible. Stimuli appeared at a randomized time between 5 s and 15 s after the start of a trial or after the needle was already inserted further than 10 % of the planned insertion depth. They were saliently rendered left to the navigation aid in blue color on a white background. The set of displayed characters was derived from the work of Weiss et al. [47]. All stimuli were rotated by 120° in both directions to ensure similar task difficulties. Fig. 4 shows the resulting set of characters. The ”4” was used for instructions only.

2.4 Sample Design

The experiment's tasks did not require specific clinical experience. Users proficient with conventional instrument navigation may even be biased towards these systems because of prior knowledge. However, a general medical background may be beneficial to understand the study's motivation and instructions and may lead to more profound participant feedback. Therefore, medical students were recruited for the user study.

2.5 Variables

The comparative user study was planned as a within-subject design single factor test. That factor was derived from the independent variable

Fig. 4. Alphanumeric characters and orientations used for secondary task. Characters were rotated 120° in both directions and mirror-reversed vertically.

display technique and consisted of four levels defined by the implemented navigation prototypes.

User performance of the primary task was measured by three dependent variables. First, the *task completion time* was ascertained. An examiner started and stopped time measurements, after participants signalized their readiness to begin and to have finished the insertion tasks respectively. Insertion accuracy was measured in two separate dimensions: *depth deviation* and *angular deviation*. Errors were analyzed independently to more specifically investigate effects of different display techniques on user performance. Depth deviation described how well participants followed the navigation visualizations insertion depth indicator. The variable was measured by calculating the absolute value of the difference between the optimal and current needle insertion depths. Values were recorded at the same time as time measurements stopped. Angular deviation measured how accurately participants adhered to the visualized insertion angle during each trial. Every 1 mm that the needle was inserted, the angle between the instrument main axis and the line connecting the actual insertion point and the target point was recorded. At the end of a trial, these data points were averaged. Values sampled at the first 5 % of the planned insertion depth were excluded from this calculation to avoid effects of large instrument movement during the initial finding of the correct insertion angle. All accuracy related calculations were based on tracking data only to reduce bias caused by prototype-dependent registration accuracy. For the same reason, deviations between planned and actual insertion positions were not investigated.

For the secondary task, a *reaction time* was measured, that was considered the time span between the appearance of the mental rotation stimulus and the time of the user-performed button press on the remote control. Consciously, only reaction times of trials with characters identified as normal (i.e. not mirror-reversed) were recorded, because reaction times reportedly differ between normal and mirror-reversed stimuli [9]. Additionally, in accordance with the literature, only correct responses were used for subsequent analyses [47]. Finally, false answers were accumulated to an *error count*.

2.6 Procedure

At the beginning of the experiment, participants' demographic data were recorded and they were instructed regarding the primary needle insertion task, the secondary mental rotation task and the study apparatus. Following, the first display technique was randomly selected and a training phase was commenced. There, detailed explanations about the active display approach were given (e.g. where the alphanumeric character will appear) and navigated needle insertions could be tested until participants felt confident to proceed. Then, three consecutive needle insertions were performed during which needed data was collected. Each insertion was accompanied with a mental rotation task. A randomized number between one and three of these tasks presented a normal (i.e. not mirror-reversed) character. Each distinct stimulus was shown no more than once. Individual trials were repeated, if the planned insertion position was missed by more than 15 mm because the extended insertion depth would otherwise no longer be comparable. A new character was randomly selected for repeated trials. After completing three proper needle insertions, the same procedure was repeated for the remaining three display techniques in randomized order. Only needed hardware was present during each trial (e.g. the tablet PC was removed if not in use). The experiment was concluded with a semi-structured interview to obtain subjective user feedback. The inquiry included the following questions:

- Do you have any general comments or questions about the study, the task, or the navigation aid?
- Do you have general comments or questions about the display devices?
- Were there one or more display devices that you liked or disliked for this task?
- Did you notice anything positive or negative when you performed the task with the monitor?
- Did you notice anything positive or negative when you performed the task with the tablet?
- Did you notice anything positive or negative when you performed the task with the AR glasses?
- Did you notice anything positive or negative when you performed the task with the projector?
- Do you have any other questions or comments?

Questions related to specific devices were asked in the order in which the display techniques were evaluated. When appropriate, responses were followed up by specific inquiries.

3 RESULTS

This section presents the data and results of the conducted user study.

3.1 Participants

Twenty-one medical students (13 female) were recruited for the user study. Participants were 20 to 35 years old (median: 25 years) and had one to seven years of university experience (median: 5 years). None of the subjects reported any degree of color vision deficiency.

3.2 Data Preparation

After completing the study, mean values of dependent variable data were calculated with respect to separate display techniques for each participant individually, to counterbalance effects of repeated measures. Regarding the reaction time variable, only correct answers of normal stimuli were considered. This resulted in missing reaction time data for one experimental condition of three participants. Measured values of these participants were excluded from subsequent analyses regarding this variable because of elsewise incomplete data rows.

3.3 Statistical Analyses

This works null hypothesis were, that on average, there is no difference between investigated display techniques regarding: task completion time ($H1_0$), depth deviation ($H2_0$), angular deviation ($H3_0$), reaction time ($H4_0$) and error count ($H5_0$). Statistical tests were set up to establish the two-sided alternative hypotheses, that on average, the selection of display techniques has different effects on respective dependent variables ($H1_a$ - $H5_a$).

Except for the error count, one-way within-subjects ANOVAs were conducted on the resulting data to investigate effects of the display techniques factor. A χ^2 goodness of fit test was applied to the accumulated error count variable. Results of the statistical analyses are summarized in Table 1. Fig. 5 visualizes identified effects.

3.4 Interpretation of Results

Statistically significant effects were found regarding task completion time and angular deviation. Both variables indicate a similar ranking between the display techniques. Thus, $H1_0$ and $H3_0$ are rejected in favor of $H1_a$ and $H3_a$, respectively. Using the projector-based AR approach required less task completion time and enabled the insertion of needles at more consistently good angles along the access path. The OST HMD method ranked second best for both variables. The monitor-based baseline condition shows the highest angular deviation. Using the VST stationary display resulted in similar high needle orientation errors and required the most amount of task completion time.

No statistically significant effects could be shown for the other variables. Therefore, $H2_0$, $H4_0$ and $H5_0$ could not be rejected. However, trends in the descriptive data suggest benefits of the projector-based approach regarding depth deviation. Especially the VST AR condition shows a high amount of variance which may have influenced the respective ANOVA's results. For the secondary task, similar reaction times were achieved for all display techniques except for the OST HMD, for which users needed more time. The highest amount of errors were made in the baseline condition. This may indicate higher mental demand for these two experimental conditions. However, too few data was collected to draw reliable conclusions on these variables.

Potential explanations to the aforementioned observations may be found in the post-hoc interview results summarized in Table 2. Similar responses were clustered and only statements with at least two mentions are reported. In addition to these results, most participants reported to have liked the projector-camera-system best for the needle

Table 1. Summary of the ANOVAs' results ($\alpha < .05$) and χ^2 goodness of fit test results on error count.

Variable	df	F	χ^2	p	Sig	η^2	Effect	Figure
Task completion time	3	13.86	-	<0.001	*	0.068	Medium	Figure 5a
Depth deviation	3	0.43	-	0.732		0.012	Small	Figure 5b
Angular deviation	3	3.29	-	0.027	*	0.075	Medium	Figure 5c
Reaction time	3	0.37	-	0.772		0.016	Small	Figure 5d
Error count	3	-	2	0.572		-	-	Figure 5e

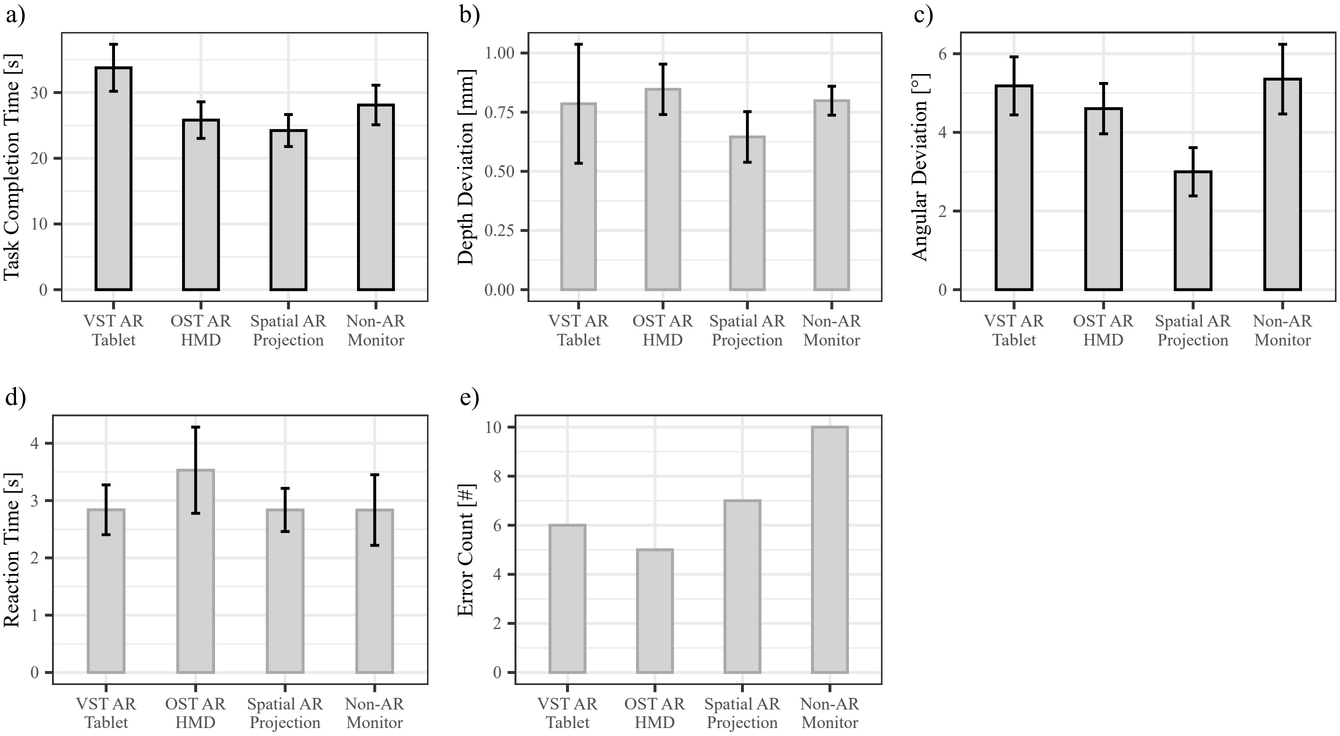

Fig. 5. Effects of display techniques on: a) task completion time, b) depth deviation, c) angular deviation, d) reaction time, and e) error count. Error bars represent standard error. Statistical significance is represented by black bar outlines. (VST - Video see-through, OST - Optical see-through)

insertion task. Some subjects rated the OST HMD similar high. Only a few users preferenced the monitor-based approach, which was also emphasized as negative by a significant amount of participants. The VST stationary display method was rather disliked by most participants. These subjective rankings are consistent with the objective data and can also be explained by the interview results.

Participants were pleased with the directly augmented view on the insertion site provided by the projector-camera-system and the OST HMD, which, at least for the projector-based AR approach, resulted in the facilitation of hand-eye-coordination. This may have also affected user performance in these conditions. For the monitor-based method, the spatial separation of navigation information and the actual injection site was experienced as demanding. Expressed difficulties in correctly placing the needle and in hand-eye-coordination further indicate this impression and may explain higher task completion times and greater angular deviations. The stationary VST display has restricted the users' scope of mobility, which resulted in rare occasions of participants switching between hands to hold the instrument and may have thus contributed to the methods extended task completion time. Additionally, the users' perspective was not in accordance with the camera stream displayed on the tablet. Depth perception was also complicated due to the reduction of the otherwise perceived three-dimensional world on a two-dimensional screen. Thus, a directly augmented view on the insertion site, like achieved with the OST HMD or the projector-camera-system, may not have been accomplished for this display technique resulting in comparably worse results.

4 DISCUSSION

The subjective participant feedback revealed aspects of the study design, that may have influenced the experiment's outcomes. The spatially fixed position of the tablet computer partially restricted participant performance in the VST AR condition. This may have been avoided by allowing position adjustments during trials, which was ultimately decided against because of additionally induced workload being only present during this method. Moreover, the rationale for a constant display position was to create comparable experimental conditions between participants, which may have been compromised by different tablet computer postures.

The OST HMD approach was affected by the goggles's small field of view and contrast capabilities. The device selection was based on hardware used in related work and a user study comparing different OST HMDs [41]. However, different devices, e.g. the second generation of the Microsoft HoloLens, may have had provided better technical capabilities. Future work should, thus, investigate effects of different OST HMDs on guided needle insertion. Moreover, focal problems of the OST-HMD in use were reported in related work [20]. The issue is inherent to the HoloLens being designed for virtual content to be placed at 2 m distance. Closer objects are perceived in wrong focal plane which may cause perceptual issues. However, no participant reported any problems regarding this topic in our experiment. Yet, the

Table 2. Summary of post-hoc interview results.

	Advantages	Disadvantages
General Comments	• Navigation aids were generally useful	• Secondary task was difficult • Using the optical tracking system was distracting
VST AR: Tablet	• Good color contrast and visualization sharpness • The video feed was useful	• Hand was obscuring the insertion site on the video • Fixed tablet position was restricting user performance and view on the phantom • Users' perspective not in accordance with the tablet • Noticeable latency • AR visualization was obscuring the needle tip • Depth perception was difficult
OST AR: HMD	• Combined view on AR visualization and phantom was useful • Good image quality • Wearing the HMD was not disturbing	• Sharpness of AR content was too low • Color contrasts were difficult to perceive • Needed head position was exhausting • The HMD's field of view was too small • AR visualization was obscuring the needle tip
Spatial AR: Projector	• Combined view on AR visualization and phantom was useful • High confidence in performance • AR Visualization facilitated hand-eye-coordination	• Projections were distorted due to surface deformations • Worse color contrast and visualization sharpness • Shadow of hand was occluding the projection
Non-AR: Monitor	• No hardware obstructed needle insertion process • Good color contrast and visualization sharpness • Looking at the monitor only did not influence performance	• Hand-eye-coordination was difficult • Placing the needle at the insertion site was difficult • Spatial separation of navigation aid and insertion site was demanding

issue may have been present and have, thus, unknowingly influenced the participants' performance outcomes.

Participants also reported problems of the shadow of their hand occluding the navigation visualization when using the projector-camera-system. This problem is commonly associated with these systems and could be mitigated by including additional projectors. However, such multi-projector-systems need to be carefully calibrated to avoid artifacts caused by unaligned projections. Additionally, surface deformations caused by inserted needles would amplify this problem and have a greater image distortion effect. Therefore, only one projector was used in this study. Besides, it is believed, that partial visualization occlusions only had a marginal effect on user performance, that could be quickly resolved by repositioning the hand. Nonetheless, potential issues in future research could be mitigated by considering shadowless projection systems as presented by Hiratani et al. [23], that require only one projector.

For both the OST HMD and the projector-camera-system impaired color perception was reported. The same color scale was employed for each device. However, different color reproduction was not regarded for the study. Because colors were used to indicate accuracy levels, this may have affected the results. Hence, more research is needed to identify suitable device-specific color scales. Moreover, effects of color correction algorithms could be explored. In the literature, such approaches were discussed for OST HMDs [25, 29] and projection systems [16, 24].

Furthermore, each implemented display technique exhibited different registration accuracy. This may have also influenced study outcomes. The effect was presumably greatest on needle positioning accuracy, which was not measured in this study. Yet, deviations from the planned insertion position prolonged the overall distance to the virtual target and, therefore, the insertion depth. This may have influenced task completion time data. However, this potential effect is believed to be marginal because of trial repetitions in case of too large positioning deviation.

Positioning deviations may also be influenced by device-related causes other than registration accuracy. For instance, participants reported problems finding the correct insertion position using the VST AR approach because of complicated depth perception. Similar effects would be worth of being investigated to determine which display technique is the most suitable to indicate the planned insertion site. Yet, such an analysis would require comparable registration results, which should be accomplished in future research.

Likewise, a future analysis of absolute targeting accuracy may yield important research findings. In this work, this accuracy measure was rejected in favor of separate angular and depth deviation parameters, because of the following rationale: Participants were instructed to maintain a stable and correct insertion angle during each needle insertion. Nonetheless, the implemented study apparatus, i.e. needle and phantom, allowed for only partially restricted angular adjustments throughout the task. This resulted in the possibility to reach the target position even if the angle was wrong for most of the time, by just applying more force and rotating the needle around the handle position rather than the surface intersection site, thus, shifting the needle away from the insertion site and partially destroying the candle gel of the phantom. Such corrections would not be possible in real interventions, where flexible needles would just bend or surgical targets would be pushed aside. Therefore, the insertion angle over time and the set insertion depth seemed to be more expressive measures to assess user errors for our study. Using such flexible needles and more realistic targets, future research may obtain expressive absolute targeting accuracy results. Resulting complicated needle tip tracking could be solved by implementing similar sensors as described by Lin et al. [31].

Hald et al. [17] conducted a similar experiment, in which a comparable set of display devices have been compared in a general sub-surface pointing task. Compared to this work's experiment, only low-depth insertions were required and targets were displayed as red circles. Hence, no navigation visualization was used. Yet, Hald et al. [17] could report similar results in terms of the projector-based method achieving best task completion time results and being the subjectively preferred method. This indicates that our findings are in accordance with the literature, and may, to some extent, be transferred to other pointing task-related research areas outside of the medical domain. However, this would require extended research.

No significant results could be shown regarding the secondary task. Participants generally commented on this task to be difficult. This may indicate, that it was, unlike intended, not completed while simultaneously inserting the needle. Instead, participants may have interrupted the primary task in favor of observing the mental rotation stimulus, thus resulting in similar reaction times and error rates. The selected character rotation angle may have also been too difficult resulting in more attention needed. Likewise, the timing of stimulus appearance was chosen, so that participants would have to perform the secondary task during the process of finding the insertion position and angle. This sub-task was believed to be the most mentally demanding. Presenting the characters at later times may have yielded different results. Moreover, more expressive results could have been obtained by an increased number of repetitions. Therefore, future research should investigate different means of assessing mental workload during guided needle insertions.

Table 2 conveys informative insights into subjective participant feedback. We decided against using standardized, quantitative questionnaires, like the system usability scale [6], in favor of conducting a semi-structured interview. We chose this method because we wanted to investigate the specific issues that participants experienced with the different methods, rather than condensing the perceived usability into a number. According to Faulkner [12], 15 participants are enough to reproducibly identify most flaws in a user interface. Specific wordings for clusters and labels might differ, if the analysis was conducted by a different team, but we believe that the general findings from this activity are reproducible and expressive and, thus, are a meaningful scientific contribution to this work.

A two-dimensional navigation aid was selected to visualize planned insertion data. Related work also implemented different techniques, which may have resulted in varying outcomes. For example, three-dimensional access paths have been displayed, on which the needles needed to be aligned [15, 20]. Especially devices with stereoscopic displaying capabilities (e.g. the used OST HMD or stereoscopic projectors) may benefit more from these navigation concepts than from the implemented one. However, three-dimensional visualizations may not be suitable for monoscopic displays (e.g. the tablet computer or the projector-camera-system), which is why a flat concept was selected for this study.

Handheld devices were excluded for this work due to the additional occupation of one hand being considered too restricting compared to the investigated display techniques. Yet, such an approach may have been advantageous compared to the stationary VST implementation because of freely adjustable display positions. Hecht et al. [18] presented a smartphone-based handheld navigation solution and revealed favorable results of their technique compared to a standard CT-guided method. Future studies may thus benefit from examining similar approaches.

Herrlich et al. investigated different alternatives to the conventional monitor-based approach. One approach mounted a small display directly on the instrument and was used to display a similar navigation aid as implemented in this work [22]. Another work used a flexible thin display that could be attached to the patient's body [21]. Both methods were shown to improve task load and usability compared to a standard monitor-based approach. Although these techniques may not be considered as augmented reality, they describe navigation solutions of interest worth analyzing in future comparisons.

Additionally, only visual displays were considered in this work. However, related work also described the implementation of auditory displays to support medical needle insertions [3,4]. As these could also be considered AR display techniques, future work may also benefit from investigating differences between such methods and implemented approaches.

5 CONCLUSION

This work attempted to close a gap in the research domain of AR supported medical needle placement, by investigating the performance of different AR display techniques and a clinically used monitor-based approach in a simulated needle insertion task for the first time. Navigation visualizations were implemented on a VST stationary display, an OST HMD and a spatial AR projector-camera-system. Performance was determined by task completion time, insertion accuracy, mental demand and subjective feedback.

Results of the user study indicate advantages of directly augmenting the view on the insertion site with navigation aids in contrast to requiring the intermediate view on screens. Especially the projector-based AR approach yielded the least amount of task completion time and angular deviation. According to subjective feedback, hand-eye-coordination would especially benefit from employing suitable AR displays. Monoscopic VST displays are not recommended because of induced depth perception problems.

In conclusion, advantages of projector-based and OST HMD AR navigation systems compared to monitor-based methods could be shown and related open research questions were determined. Results contribute to the development and design of improved needle navigation systems. Future work should focus on improving favored investigated methods regarding issues identified in this study.

ACKNOWLEDGMENTS

This work was funded by the German Research Foundation (HA 7819/1-2 & LA 3855/1-2).

REFERENCES

[1] C. A. Agten, C. Dennler, A. B. Rosskopf, L. Jaberg, C. W. Pfirrmann, and M. Farshad. Augmented Reality-Guided Lumbar Facet Joint Injections. *Investigative Radiology*, 53(8):495–498, 2018.

[2] G. Ballestin, F. Solari, and M. Chessa. Perception and action in peripersonal space: A comparison between video and optical see-through augmented reality devices. In *Proc. of International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 184–189. IEEE, 2018.

[3] D. Black, J. Hettig, M. Luz, C. Hansen, R. Kikinis, and H. Hahn. Auditory feedback to support image-guided medical needle placement. *International Journal of Computer Assisted Radiology and Surgery*, 12(9):1655–1663, 2017.

[4] F. Bork, B. Fuers, A.-K. Schneider, F. Pinto, C. Graumann, and N. Navab. Auditory and visio-temporal distance coding for 3-dimensional perception in medical augmented reality. In *Proc. of International Symposium on Mixed and Augmented Reality*, pp. 7–12. IEEE, 2015.

[5] P. Breedveld, H. Stassen, D. Meijer, and L. Stassen. Theoretical background and conceptual solution for depth perception and eye-hand coordination problems in laparoscopic surgery. *Minimally Invasive Therapy & Allied Technologies*, 8(4):227–234, 1999.

[6] J. Brooke. Sus: a "quick and dirty'usability. *Usability evaluation in industry*, p. 189, 1996.

[7] J. F. Buell, M. T. Thomas, S. Rudich, M. Marvin, R. Nagubandi, K. V. Ravindra, G. Brock, and K. M. McMasters. Experience with more than 500 minimally invasive hepatic procedures. *Annals of surgery*, 248(3):475–486, 2008.

[8] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, and M. Ivkovic. Augmented reality technologies, systems and applications. *Multimedia Tools and Applications*, 51(1):341–377, 2011.

[9] L. A. Cooper and R. N. Shepard. The time required to prepare for a rotated stimulus. *Memory & Cognition*, 1(3):246–250, 1973.

[10] M. Das, F. Sauer, U. J. Schoepf, A. Khamene, S. K. Vogt, S. Schaller, R. Kikinis, E. vanSonnenberg, and S. G. Silverman. Augmented reality visualization for CT-guided interventions: system description, feasibility, and initial evaluation in an abdominal phantom. *Radiology*, 240(1):230–235, 2006.

[11] Z. Fan, L. Ma, Z. Liao, X. Zhang, and H. Liao. Three-dimensional image-guided techniques for minimally invasive surgery. In *Handbook of Robotic and Image-Guided Surgery*, pp. 575–584. Elsevier, 2020.

[12] L. Faulkner. Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, 35(3):379–383, 2003.

[13] J. Fritz, P. U-Thainual, T. Ungi, A. J. Flammang, N. B. Cho, G. Fichtinger, I. I. Iordachita, and J. A. Carrino. Augmented reality visualization with image overlay for MRI-guided intervention: accuracy for lumbar spinal procedures with a 1.5-T MRI system. *American Journal of Roentgenology*, 198(3):W266–W273, 2012.

[14] K. Gavaghan, T. Oliveira-Santos, M. Peterhans, M. Reyes, H. Kim, S. Anderegg, and S. Weber. Evaluation of a portable image overlay projector for the visualisation of surgical navigation data: phantom studies. *International Journal of Computer Assisted Radiology and Surgery*, 7(4):547–556, 2012.

[15] J. T. Gibby, S. A. Swenson, S. Cvetko, R. Rao, and R. Javan. Head-mounted display augmented reality to guide pedicle screw placement utilizing computed tomography. *International Journal of Computer Assisted Radiology and Surgery*, 14(3):525–535, 2019.

[16] A. Grundhöfer and D. Iwai. Robust, error-tolerant photometric projector compensation. *IEEE Transactions on Image Processing*, 24(12):5086–5099, 2015.

[17] K. Hald, M. Rehm, and T. B. Moeslund. Augmented reality technology for displaying close-proximity sub-surface positions. In *Proc of. IFIP Conference on Human-Computer Interaction*, pp. 641–659. Springer, 2019.

[18] R. Hecht, M. Li, Q. M. de Ruiter, W. F. Pritchard, X. Li, V. Krishnasamy, W. Saad, J. W. Karanian, and B. J. Wood. Smartphone augmented reality CT-based platform for needle insertion guidance: A phantom study. *CardioVascular and Interventional Radiology*, pp. 1–9, 2020.

[19] F. Heinrich, F. Joeres, K. Lawonn, and C. Hansen. Comparison of projective augmented reality concepts to support medical needle insertion. *IEEE Transactions on Visualization and Computer Graphics*, 25(6):2157–2167, 2019.

[20] F. Heinrich, L. Schwenderling, M. Becker, M. Skalej, and C. Hansen. HoloInjection: augmented reality support for CT-guided spinal needle injections. *Healthcare Technology Letters*, 6(6):165–171, 2019.

[21] M. Herrlich, A. V. Reinschluessel, M. Willems, N. Langhorst, D. Black, T. Doering, C. Rieder, R. Kikinis, and R. Malaka. Put that needle there: Customized flexible on-body thin-film displays for medical navigation. *ACM Transactions on Computing for Healthcare*, 2020.

[22] M. Herrlich, P. Tavakol, D. Black, D. Wenig, C. Rieder, R. Malaka, and R. Kikinis. Instrument-mounted displays for reducing cognitive load during surgical navigation. *International Journal of Computer Assisted Radiology and Surgery*, 12(9):1599–1605, 2017.

[23] K. Hiratani, D. Lwai, P. Punpongsanon, and K. Sato. Shadowless projector: Suppressing shadows in projection mapping with micro mirror array plate. In *Proc. of IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 1309–1310. IEEE, 2019.

[24] B. Huang and H. Ling. End-to-end projector photometric compensation. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6810–6819, 2019.

[25] Y. Itoh, M. Dzitsiuk, T. Amano, and G. Klinker. Semi-parametric color reproduction method for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1269–1278, 2015.

[26] M. C. Juan and J. Calatrava. An augmented reality system for the treatment of phobia to small animals viewed via an optical see-through HMD: Comparison with a similar system viewed via a video see-through HMD. *International Journal of Human-Computer Interaction*, 27(5):436–449, 2011.

[27] R. Krempien, H. Hoppe, L. Kahrs, S. Daeuber, O. Schorr, G. Eggers, M. Bischof, M. W. Munter, J. Debus, and W. Harms. Projector-based augmented reality for intuitive intraoperative guidance in image-guided 3D interstitial brachytherapy. *International Journal of Radiation Oncology, Biology, Physics*, 70(3):944–952, 2008.

[28] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *Proc. of International Symposium on Mixed and Augmented Reality*, pp. 3–12. IEEE, 2010.

[29] T. Langlotz, M. Cook, and H. Regenbrecht. Real-time radiometric compensation for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2385–2394, 2016.

[30] J. S. Lewin, C. Thomas, P. L. Pereira, S. Clasen, C. D. Claussen, and J. Fritz. Freehand Real-Time MRI-Guided Lumbar Spinal Injection Procedures at 1.5 T: Feasibility, Accuracy, and Safety. *American Journal of Roentgenology*, 192(4):W161–W167, 2009.

[31] M. A. Lin, A. F. Siu, J. H. Bae, M. R. Cutkosky, and B. L. Daniel. Holoneedle: Augmented reality guidance system for needle placement investigating the advantages of three-dimensional needle shape reconstruction. *IEEE Robotics and Automation Letters*, 3(4):4156–4162, 2018.

[32] M. Luz, G. Strauss, and D. Manzey. Impact of image-guided surgery on surgeons' performance: a literature review. *International Journal of Human Factors and Ergonomics*, 4(3-4):229–263, 2016.

[33] D. Manzey, S. Röttger, J. E. Bahner-Heyne, D. Schulze-Kissing, A. Dietz, J. Meixensberger, and G. Strauss. Image-guided navigation: the surgeon's perspective on performance consequences and human factors issues. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 5(3):297–308, 2009.

[34] A. Marien, A. C. de Luis Abreu, M. Desai, R. A. Azhar, S. Chopra, S. Shoji, T. Matsugasumi, M. Nakamoto, I. S. Gill, and O. Ukimura. Three-dimensional navigation system integrating position-tracking technology with a movable tablet display for percutaneous targeting. *BJU international*, 115(4):659–665, 2015.

[35] A. Mewes, F. Heinrich, U. Kägebein, B. Hensen, F. Wacker, and C. Hansen. Projector-based augmented reality system for interventional visualization inside MRI scanners. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 15(1):e1950, 2019.

[36] C. Moser, J. Becker, M. Deli, M. Busch, M. Boehme, and D. H. Groenemeyer. A novel Laser Navigation System reduces radiation exposure and improves accuracy and workflow of CT-guided spinal interventions: A prospective, randomized, controlled, clinical trial in comparison to conventional freehand puncture. *European Journal of Radiology*, 82(4):627–632, 2013.

[37] T. Munzner. *Visualization analysis and design*. AK Peters/CRC Press, 2014.

[38] A. Nee, S. Ong, G. Chryssolouris, and D. Mourtzis. Augmented reality applications in design and manufacturing. *CIRP Annals*, 61(2):657 – 679, 2012.

[39] J.-M. Normand, M. Servières, and G. Moreau. A new typology of augmented reality applications. In *Proc. of the 3rd Augmented Human International Conference*, AH '12. Association for Computing Machinery, New York, NY, USA, 2012.

[40] T. Oliveira-Santos, B. Klaeser, T. Weitzel, T. Krause, L. P. Nolte, M. Peterhans, and S. Weber. A navigation system for percutaneous needle interventions based on PET/CT images: design, workflow and error analysis of soft tissue and bone punctures. *Computer aided surgery*, 16(5):203–219, 2011.

[41] L. Qian, A. Barthel, A. Johnson, G. Osgood, P. Kazanzides, N. Navab, and B. Fuerst. Comparison of optical see-through head-mounted displays for surgical interventions with object-anchored 2D-display. *International Journal of Computer Assisted Radiology and Surgery*, 12(6):901–910, 2017.

[42] J. P. Rolland and H. Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoperators & Virtual Environments*, 9(3):287–309, 2000.

[43] P. Schullian, G. Widmann, T. B. Lang, M. Knoflach, and R. Bale. Accuracy and diagnostic yield of CT-guided stereotactic liver biopsy of primary and secondary liver tumors. *Computer aided surgery*, 16(4):181–187, 2011.

[44] A. Seitel, N. Bellemann, M. Hafezi, A. M. Franz, M. Servatius, A. Saffari, T. Kilgus, H.-P. Schlemmer, A. Mehrabi, B. A. Radeleff, and L. Maier-Hein. Towards markerless navigation for percutaneous needle insertions. *International Journal of Computer Assisted Radiology and Surgery*, 11(1):107–117, 2016.

[45] R. Vassallo, A. Rankin, E. C. Chen, and T. M. Peters. Hologram stability evaluation for microsoft hololens. In *Proc. of Medical Imaging: Image Perception, Observer Performance, and Technology Assessment*, vol. 10136, p. 1013614. International Society for Optics and Photonics, 2017.

[46] F. K. Wacker, S. Vogt, A. Khamene, J. A. Jesberger, S. G. Nour, D. R. Elgort, F. Sauer, J. L. Duerk, and J. S. Lewin. An augmented reality system for MR image–guided needle biopsy: initial results in a swine model. *Radiology*, 238(2):497–504, 2006.

[47] M. M. Weiss, T. Wolbers, M. Peller, K. Witt, L. Marshall, C. Buchel, and H. R. Siebner. Rotated alphanumeric characters do not automatically activate frontoparietal areas subserving mental rotation. *Neuroimage*, 44(3):1063–1073, 2009.