**Florian Heinrich, Fabian Joeres,**

**Kai Lawonn, Christian Hansen**

# Comparison of Projective Augmented Reality Concepts to Support Medical Needle Insertion

Pre-print version

**Florian Heinrich, Fabian Joeres,**

**Christian Hansen**

Faculty of Computer Science,

Otto-von-Guericke University Magdeburg, Germany

hansen@isg.cs.uni-magdeburg.de


**Kai Lawonn**

Faculty of Computer Science,

University of Koblenz-Landau, Germany

lawonn@uni-koblenz.de

# Comparison of Projective Augmented Reality Concepts to Support Medical Needle Insertion

Florian Heinrich, Fabian Joeres, Kai Lawonn, Christian Hansen

**Abstract**—Augmented reality (AR) is a promising tool to improve instrument navigation in needle-based interventions. Limited research has been conducted regarding suitable navigation visualizations. In this work, three navigation concepts based on existing approaches were compared in a user study using a projective AR setup. Each concept was implemented with three different scales for accuracy-to-color mapping and two methods of navigation indicator scaling. Participants were asked to perform simulated needle insertion tasks with each of the resulting 18 prototypes. Insertion angle and insertion depth accuracies were measured and analyzed, as well as task completion time and participants' subjectively perceived task difficulty. Results show a clear ranking of visualization concepts across variables. Less consistent results were obtained for the color and indicator scaling factors. Results suggest that logarithmic indicator scaling achieved better accuracy, but participants perceived it to be more difficult than linear scaling. With specific results for angle and depth accuracy, our study contributes to the future composition of improved navigation support and systems for precise needle insertion or similar applications.

**Index Terms**—Visualization, augmented reality, evaluation, medical navigation systems, instrument guidance, needle placement.

---

## 1 INTRODUCTION

THE proven benefits of minimally invasive surgery in terms of lower risk of infections and damage to healthy tissue led to a significant trend away from open surgery and towards minimally invasive interventions in recent years [1]. Treatments like tumor ablations and biopsies usually require a precisely targeted insertion of needle-shaped instruments. During such procedures, the patient is not opened, which means that visual and haptic feedback of anatomical and pathological structures as well as surgical instruments are missing. During interventions, radiological imaging (i.e. ultrasound, computed tomography (CT) or magnetic resonance imaging (MRI)) is essential to compensate for this problem and to successfully locate target regions [2], [3].

To further improve the quality of interventions, surgical navigation systems have been developed. Using these systems can decrease the general procedure time, the number of required imaging scans and improve the targeting accuracy [4], [5]. Thereby, inserted instruments are usually tracked and their position is visualized together with anatomical images and additional navigation information, like the distance to the target [6].

However, such information and images are usually presented on an external monitor, spatially disjunct from the operation site. This separation of useful information and the patient increases mental load and time pressure [7]. Consulting the navigation display also interrupts the interventionalist's attention to the patient and leads to complicated hand-eye coordination [8]. Visualizing instrument guidance information directly on the patient in the form of augmented reality (AR) can help with these issues [9]. Previous advances in this field yielded feasible results and made use of different display modalities, like video see-through monitors [10] and optical see-through head-mounted displays (HMDs) [11]. Gavaghan et al. proposed a projective AR approach in this field and concluded that their method may overcome deficiencies of previous advances [9].

While much research has been conducted in the domain of developing AR instrument navigation systems as a whole, fewer publications focus on questions regarding the design of navigation visualizations. In this work, we compared three visualization concepts derived from an analysis of existing approaches. To determine the most applicable method in terms of targeting accuracy, we conducted a user study and thereby considered different parameters in accuracy-to-color mapping and navigation indicator scaling. To focus on visualization effectiveness, we reduced the effects of system calibration accuracy by measuring the accuracy with which users followed the visualizations' guidance, rather than measuring absolute accuracy in real needle position. We displayed the navigation visualizations in a projective AR setup, as it is a promising display modality approach. A supplemented video gives a summary of evaluated visualizations and the user study.

## 2 RELATED WORK

Most related work focuses on the development of instrument guidance systems and/or the evaluation of resulting systems. A significant share of visualization concepts are described in navigation system feasibility studies with the primary objective of measuring absolute insertion accuracy. These studies are often published in medical journals which focus less on technical aspects. An overview of these systems' navigation methods is given in the following.

- F. Heinrich, F. Joeres and C. Hansen are with University of Magdeburg, Germany, and Research Campus STIMULATE.
  E-mail: hansen@isg.cs.uni-magdeburg.de
- K. Lawonn is with University of Koblenz-Landau, Germany.
  E-mail: lawonn@uni-koblenz.de

## 2.1 See-through Vision Concepts

Several groups propose supporting needle placement by providing some sort of see-through vision into the patient's body. These concepts visualize the target structure as well as an extrapolation of the instrument trajectory inside of the patient. Das et al. used a stereoscopic video see-though HMD to provide spatially correctly positioned radiological images together with the view on a virtual needle and its trajectory's extrapolation. They evaluated their method by performing simulated biopsy insertion tasks and measuring placement errors in CT images [12]. Fritz et al. followed a similar approach using an optical see-through display positioned in front of an MRI bore. The display offered a spatially correct view on a radiological image slice augmented with a pre-planned insertion path. To evaluate their method the authors had participants perform puncture tasks on a lumbar spine phantom. They used CT images to determine various dimensions of absolute accuracy errors [13].

Kanithi et al. described an optical see-through HMD approach to fuse the user's vision with an extrapolation of a needle and ultrasound images positioned directly at the ultrasound probe. The authors focused on the evaluation of system calibration accuracy and did not assess user performance data yet [14]. Bork et al. developed methods to improve spatial perception for medical AR applications by visualizing distance information between a surgical instrument and regions of interest in a superimposed video stream. To evaluate their methods they simulated needle biopsies of virtual lesions with and without their techniques and measured the distance between the tracked needle and the target positions [15].

## 2.2 Access Path Visualization Concepts

Several other groups propose methods similar to see-through vision visualizations. These concepts may visualize anatomical structures as well, but focus more on the visualization of the planned access path. For such methods, the needle needs to be aligned with the emphasized path. Seitel et al. followed this approach in a monitor-based setup. Desired needle positions are displayed in a camera view on the patient and the needle has to be held so as to be completely occluded by its virtual representation. The authors evaluated both user error (i.e. distance between tracked needle tip and target) and absolute targeting error of needle insertion tasks (i.e. distance between needle tip and target in CT images) [10]. Kuzhagaliyev et al. used an optical see-through HMD to display the access path as a colored line and displayed the actual needle path together with a triangle indicating the deviation between these trajectories. However, the approach was not yet evaluated regarding insertion accuracy [16].

In a monitor-based approach, Khlebnikov et al. visualized possible access paths as crepuscular rays pointed towards the target region and colored according to risk assessment calculations. The visualization was not evaluated for insertion tasks since it was developed for accessibility planning only [17]. Alpers et al. presented a monitor-based navigation system for tumor ablations. They guide needle insertion by visualizing the current needle pathway and anatomical risk structures along its trajectory on radiological images. The method was evaluated in a virtual insertion experiment where a needle was controlled via mouse input [18].

## 2.3 Explicit Navigation Aids Concepts

Some navigation systems provide the user with explicit navigation aids on how to position, orient and insert the instrument. Oliveira-Santos et al. developed a system which uses a crosshairs-shaped visualization to guide the insertion process. Needle tip and handle projections need to be aligned at the crosshairs' center for positioning and orientation adjustments. A progress bar represents the total insertion depth. The authors evaluated their system with a needle insertion experiment simulating lesions in swine cadavers. Absolute targeting accuracy was measured and decomposed into a user specific error, a tissue deformation error and an error resulting from patient movement [6]. A similar approach was followed by Gavaghan et al. in case of a projective AR scenario where they visualized the crosshairs at the puncture position. They evaluated their approach in a qualitative feasibility study [9].

Wacker et al. combined a see-through vision method with a depth control visualization through direct navigation cues. Two concentric circles of different diameters are positioned at the center of a target and merge as the needle gets inserted. Cadaver and phantom studies were conducted to evaluate system accuracy. Inserted needle accuracy was verified on MR images [19]. Krempien et al. proposed similar navigation cues for depth control using projective AR. Additionally they projected an arrow beginning at the insertion position and pointing towards the correct insertion angle. The approach was evaluated in a clinical study. CT scans were used to measure insertion accuracy [20].

Kreiser et al. focused on supporting the placement of multiple needles. For each needle, a navigation visualization was arranged around a circular grid. Each visualization consists of a set of circle and arrow glyphs that need to be aligned over a needle-shaped frame. The method was evaluated in a quantitative user study and compared to a standard method using angle, depth and distance error metrics [21].

## 2.4 Comparative AR Instrument Navigation Studies

A few publications explicitly focus on the design of possible navigation visualizations and the comparison of different approaches. These studies' scope and objectives are similar to this article's. Thus, the studies are briefly outlined below. Seitel et al. compared four different computer-assisted targeting visualizations. They measured tip positioning and insertion depth accuracies between a tracked instrument and planned insertion and target points. Their concepts were designed for use on an external monitor, and visualized three-dimensional (3D) scenes [22]. These virtual scenes would need to be registered to the operation site to be used with AR visualizations, which may not be applicable for all of their proposed methods. Furthermore, concepts might perform differently when used and evaluated in an AR setup.

Chan and Heng researched different methods to visualize access paths for optimal needle placements. In a user

study, they found their approach of displaying the access path as a volumetric beam together with an array of rings to be superior in terms of displacement and tilting errors of virtual needles compared to other techniques [23]. However, those access path visualizations were designed for the use on an external monitor, as well. The adaptivity of such visualizations to projective AR would have to be further investigated first.

Mewes et al. described the design of two projective AR visualization concepts for needle guidance inside an MRI. The first approach uses see-through vision and projects the target position along with the needle trajectory onto a body. As a second concept, they proposed direct navigation cues. An arrow visualization supports the needle orientation task and a progress bar represents the insertion depth. In a comparative study, Mewes et al. measured absolute target distance and positioning errors verified by MRI images [24]. However, target distance errors are composed of three separate dimensions - positioning, orientation and depth accuracy. The authors only measured the positioning accuracy, so no conclusions are possible regarding their concepts' suitability for the other two dimensions.

## 3 TECHNICAL METHODS

Several instrument navigation concepts were presented in the previous section. Not all of these concepts are suitable for our purposes, because we intend to make use of projective AR. Access path visualizations have not been used for projective AR before, and see-through vision concepts often rely on radiological image support, which may not be feasible due to projection specific drawbacks (e.g. distortion, reduced resolution and contrast). Additionally, concepts based on the visualization of 3D scenes heavily rely on correct depth perception, which is a major issue for projective AR [25]. Therefore, we decided to examine two-dimensional explicit navigation aids and focus on approaches which have already been successfully implemented for projective AR [9], [20], [24]. This section presents three investigated navigation concepts and two different means to parameterize the visualizations. Fig. 1, Fig. 2, and Fig. 3 show the functionality of the three concepts during four different insertion steps each. The images presented represent identical insertion situations

### 3.1 Circle Concept

In the first concept we adopted the idea of encoding navigation information through circular glyphs. Wacker et al. [19] and Krempien et al. [20] visualized two concentric circles where the difference between their diameters represented the current insertion depth. In a similar fashion, our method displays concentric circles to encode navigation information. Fig. 1 shows an overview of this concept's functionality.

In the beginning of an insertion process, two concentric circular arcs are positioned at a predefined maximum radius from their center point, which coincides with the planned puncture position. The further an instrument is inserted, the smaller the diameter of the inner arc becomes. The space between the arcs is filled with color. That color's hue encodes depth accuracy information. At maximum insertion depth, the inner arc completely vanishes. Inserting the
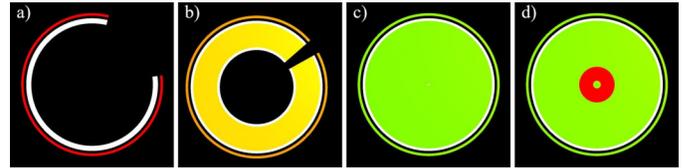


Fig. 1. Circle Concept. Concentric circular arcs encode insertion depth and angular accuracies. Arcs are opened towards planned insertion trajectory. Opening angles represent angular deviations and arc diameters encode insertion depth information. From left to right, the insertion angle is steadily improved and the needle is inserted further. (d) shows a needle that was inserted too far. Colors are mapped to accuracy levels.

needle further results in the visualization of a new circular arc of increasing diameter. The displayed circular arcs are opened towards the direction the needle has to be tilted to match the correct insertion orientation. The higher the magnitude of the angular deviation, the larger the angle of the opening in the circular arcs. A correctly oriented needle results in full 360° arcs. For further assistance, an additional concentric circular arc is displayed at the outermost radius. The arc's color encodes angular accuracy information.

### 3.2 Crosshairs Concept

Several publications describe crosshairs-shaped navigation methods [6], [9]. We adapted the idea for the implementation of our second navigation concept. A crosshairs-shaped glyph is projected onto the body with the center of the crosshairs placed at the planned puncture position. Fig. 2 shows an overview of the visualization.

The concept visualizes insertion depth information as a colored circle of varying diameter. The circle increases in size as the needle is inserted. The border of the circle reaches the outer ring of the crosshairs visualization when the needle is inserted to the maximum insertion depth. The circle diameter increases even further when the needle gets inserted too far. The hue of the circle encodes depth accuracy information. To support users to find the correct insertion angle, the position of the needle handle is projected onto the crosshairs visualization as if it was positioned on a plane perpendicular to the planned insertion trajectory. The handle position is marked by a smaller crosshairs glyph. The distance between this marker and the center of the larger crosshairs represents the magnitude of the angular deviation. Tilting the needle results in movement of the marker. Additionally, the marker is filled with color to encode angular accuracy information.
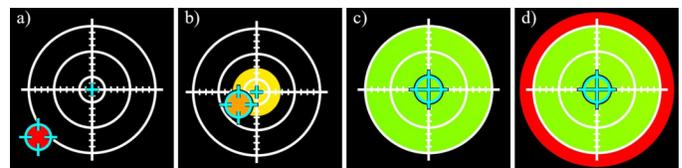


Fig. 2. Crosshairs Concept. The needle handle is projected onto a crosshairs grid (blue). For correct needle orientation, the marker has to align with the crosshairs' center. The grid is steadily filled to encode insertion depth information. From left to right the insertion angle is steadily improved and the needle is inserted further. (d) shows a needle that was inserted too far. Colors are mapped to accuracy levels.

## 3.3 Arrow Concept

As a third concept, we adopted the method of Mewes et al. [24], who used projective AR to display direct navigation cues on the patient. This concept visualizes an arrow indicating angular deviations. The arrow starts at the planned needle insertion site, thus indicating where to place the needle. Fig. 3 shows the navigation concept during different insertion states.

The concept visualizes depth information in form of a progress bar positioned to the left of the arrow indicator. The bar's filling level increases as the needle is inserted. An over-filled progress bar indicates that the maximum insertion depth has been exceeded. The color of the filling is used to encode depth accuracy information. The arrow glyph describes insertion angle information. The arrow is pointed towards the planned insertion trajectory and varies in length according to the angle deviation magnitude. A smaller arrow indicates smaller angular deviations. When the needle is oriented correctly, the arrow is replaced by a small circle. The hue of the arrow's color encodes angular accuracy information.
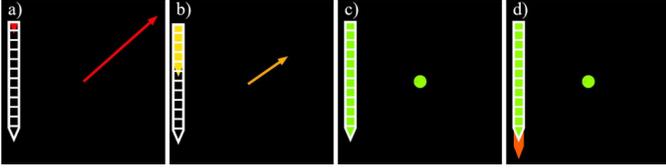


Fig. 3. Arrow Concept. An arrow glyph indicates angular deviations and a progress bar represents insertion depth information. The arrow is pointed towards the planned insertion trajectory. Its length represents angular deviations. From left to right the insertion angle is steadily improved and the needle gets further inserted. (d) shows a needle that was inserted too far. Colors are mapped to accuracy levels.

## 3.4 Accuracy-to-Color Mapping

All described concepts use color to encode angular and depth accuracies. Fig. 4 shows the implemented color scales. A traffic light metaphor was used for easier interpretation of hues. In case of study participants with color vision impairments, a different scale using blue hues was implemented, as well. We differentiated between discrete and continuous color mapping.

For discrete color mapping, four colors were assigned to different threshold levels , as indicated by Fig. 4. Values were based on clinical feedback regarding targeting accuracy needs for the ablation of tumors with a size of 15mm [26]. The same colors were applied to angular deviations. The thresholds were chosen to aid needle alignment as closely as possible up to reasonable accuracy for free-hand insertion tasks.

Continuous accuracy-to-color mapping uses the same threshold levels. The only exceptions are posed by the smallest thresholds which are set to 0mm and 0° respectively. While the discrete method only displays the colors assigned to the respective accuracy levels, the continuous mapping interpolates linearly between colors. Therefore, accuracies between thresholds are mapped to distinct colors, as well.
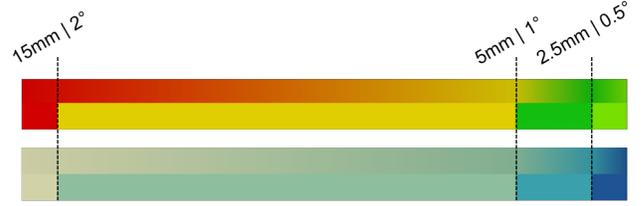


Fig. 4. Accuracy-to-color mapping scales. Color range represents poor (left) to good (right). Both continuous and discrete mapping scales were considered. The top scale uses a traffic light metaphor and the bottom scale is used in case of vision impairments. Color anchors were set to absolute accuracy values at marked positions, resulting in non-linear color distributions.

## 3.5 Indicator Scaling

The implemented concepts use individual indicators to encode angular and depth deviations as described above (e.g. the Arrow concept indicates angular deviations by the length of the arrow and depth deviations by the filling level of the progress bar). We implemented two methods to scale the movement of these indicators. A linear indicator scaling uses linear interpolation to calculate indicator movements. In case of the Arrow concept, changes of needle orientation directly translate to changes of the arrow's length. In contrast, we also implemented a different indicator scaling using the logarithmic function below:

$$f(x, b, t) = \log_{b+1}(1 + (x \cdot \sqrt[t]{b})^t), \qquad (1)$$

where x is the respective linear scaling factor between $0.0$ and $1.0$, b is the logarithmic base and t is a floating-point number amplifying or damping the logarithmic effect. The base variable b can also be varied to the manipulate the logarithmic effect. The resulting function value is then multiplied by each indicator's maximum size to calculate the scaling effect. The formula causes greater indicator movement changes at high accuracy levels and smaller movement changes at worse accuracy levels. In case of the Arrow concept needle orientation changes closer to a deviation of 0° cause a greater reduction of the arrow's length, than changes at a deviation of 20°. For our implementation, we used parameters b=100 and t=1.5. The values were selected based on preliminary study results and seem to yield the desired scaling effect.

## 3.6 Technical Realization

All described concepts have been implemented using the game engine *Unity* (Unity Technologies, USA) for its wide-ranging amount of available computer graphics functionalities. Navigation cues were implemented as procedurally generated meshes at runtime. To display the navigation concepts directly on a body, we followed a projective AR approach. In our lab setup, we mounted a Barco F22 WUXGA Digital Light Processing projector (Barco GmbH, Karlsruhe, Germany) to the top left side from the user, using an aluminum frame. The projection system can be seen in Fig. 5 and was calibrated using the photogrammetric measurement system *ProjectionTools* (domeprojections.com GmbH, Berlin, Germany). Resulting parameters were a scan of the projection surface, projector position, orientation and frustum as well as lens distortion coefficients.
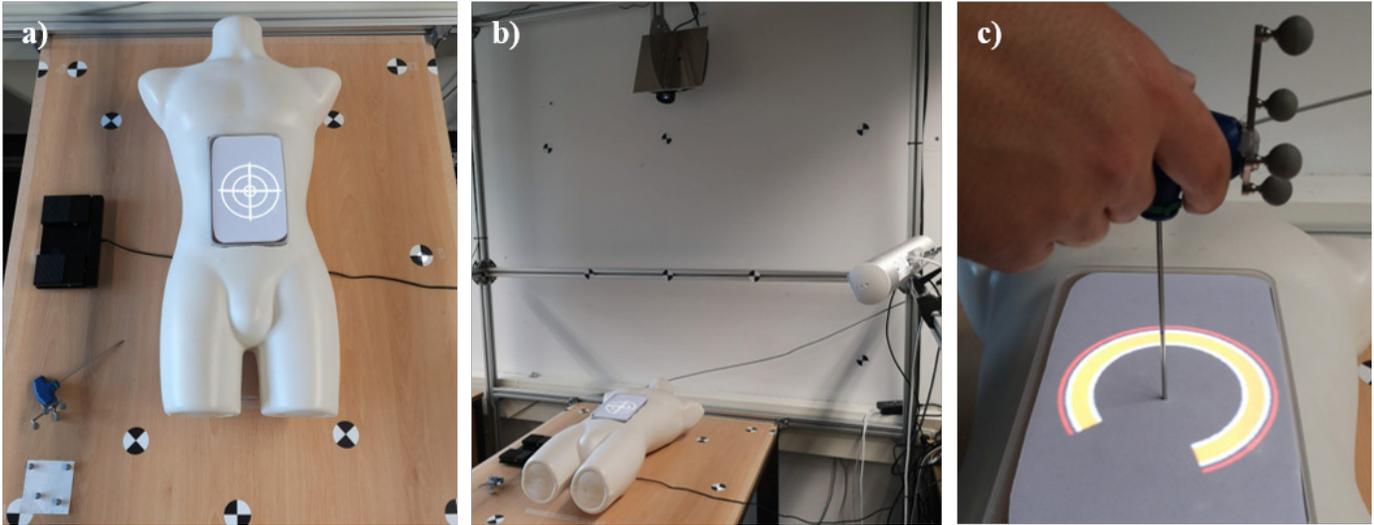
Fig. 5. The technical setup. (a) and (b) show an overview of the projection system and materials for the user study. A phantom used as projection surface is positioned on a table. A switch used for interaction is positioned beside the phantom. The bottom left corner (from the camera) of the table defined the global world coordinate system and is indicated by optical markers. A projector (b - top) is mounted above the table and an optical tracking system is oriented towards the projection surface (b - right). (c) shows an insertion process guided by the Circle concept (see Fig. 1). The user holds a needle applicator with optical markers used for tracking.

The coordinates of the projection surface and the projector position were calculated in a mutual world coordinate system defined by an optical marker. To spatially and perspectively correctly project navigation aids, the rendering was shared with a middleware responsible for calculating undistorted projection images. We used an infrared-light-based (IR) optical tracking system (fusionTrack 500, Atracsys LLC, Switzerland) to retrieve instrument position data. Calibration between projection and tracking systems was done using an optical IR marker with a known transform to the optical marker defining the world coordinate system. Fig. 5 shows the spatially correct projection of a navigation cue on a phantom as the final result of all calibration and calculation steps. We did not measure calibration accuracy data because this work only focused on visually supporting simulated needle insertions. Therefore, our purposes did not require clinical calibration accuracy at sub-millimeter level.

## 4 EVALUATION

After the implementation of navigation visualizations, a user study was conducted. The study's goal was to compare the introduced concepts and to find an optimal set of parameters to improve navigation support. This section describes the selected independent and dependent variables, participant tasks, and the experimental procedure.

### 4.1 Tasks

For each presented prototype, participants were asked to complete a needle insertion task. We simulated clinical procedures by using a phantom filled with candle gel covered by a sheet of paper to provide a haptic experience similar to skin. The participants were asked to insert a tracked needle into the candle gel while keeping a predefined insertion angle as accurately as possible. A predefined insertion depth had to be reached as closely as possible and participants were told not to over-insert the needle.

### 4.2 Sample Design

We recruited medical students for this study because our research objectives did not require the participants to have any specific clinical experience. However, a general medical background may help participants to mentally put the application into its context and understand the motivation behind the instructions (e.g. over-insertion is more critical than under-insertion). Twenty-five (25) medical students (14 female, 11 male) participated in the study. Participant age ranged from 21 years to 27 years old (median: 24 years) and participants were between their 2nd and 6th year of university. Three participants had some degree of deuteranopia.

### 4.3 Independent Variables

To compare the concepts and to investigate the effect of different parameters, a three-factor test was carried out. The three factors were defined by the independent variables *Concept*, *Color*, and *Scaling*. The variable *Concept* was defined as a factor with three levels consisting of the Circle, Crosshairs, and Arrow concepts (see Sect. 3.1 to Sect. 3.3). *Color* refers to the previously described accuracy-to-color mapping scales (see Sect. 3.4) and was defined as a factor with three levels. Besides the discrete and continuous scales, a monochromatic factor level was used to analyze the impact of disabling accuracy-to-color mapping by using a plain white color scale. The variable *Scaling* was defined as a factor with two levels and refers to the methods of linear and logarithmic indicator scaling (see Sect. 3.5).

### 4.4 Dependent Variables

During each trial, we measured five dependent variables to compare the 18 prototypes resulting from all factor level combinations. Each of the three dimensions positioning, orientation and depth need to be regarded separately to correctly measure and interpret insertion accuracy. Since

positioning accuracy was already investigated by Mewes et al. [24] and seems to be a problem related to system calibration and registration errors, we only focused on the dimensions orientation and depth accuracy. Angular and depth deviations were measured separately and did not interfere with one another. For the purpose of an individual evaluation, depth and angle navigation cues were calculated and visualized independently. We used optical tracking data to calculate virtual needle positions a representation of the physical object's positions and to calculate placement accuracy parameters.

During each trial, we measured insertion depth as the distance between the virtual needle tip and the insertion point. We subtracted the final measured insertion depth from the targeted insertion depth and interpreted the absolute value of the difference as an *absolute depth deviation*. To measure orientation accuracies, we repeatedly calculated the angle between the current needle trajectory and the trajectory of the pre-planned insertion path. After the first 10mm, a new angular deviation was calculated every $3.\bar{3}$mm along the insertion process. All data points were averaged at the end of each trial. The resulting *angular deviation* therefore describes the quality of needle orientation throughout the insertion process. Additionally, we measured the *task completion time* for each trial. Time measurement started when participants signalized their readiness by pressing a single-button switch near their standing position. Time measurement was stopped when participants pressed the button again. At that time, the final insertion depth was recorded. After the completion of each trial, participants were asked to rate perceived difficulty to find the correct angle and depth using the given navigation prototype. The dimensions *subjective difficulty angle* and *subjective difficulty depth* were rated separately on a 6-point Likert scale with verbal anchors "very easy", "easy", "slightly easy", "slightly difficult", "difficult" and "very difficult".

### 4.5 Hypotheses and statistical analysis

We applied five three-factorial Analyses of variance (ANOVA) to investigate the following two-sided alternative hypotheses:
**H1:** The *concept* factor has an effect on the five dependent variables: angular accuracy (H1.1), depth accuracy (H1.2), task completion time (H1.3), subjective rating of angle setting difficulty (H1.4), and subjective rating of depth setting difficulty (H1.5).
**H2:** The *color* factor has an effect on the five dependent variables (H2.1 - H2.5).
**H3:** The *scaling* factor has an effect on the five dependent variables (H3.1 - H3.5)

### 4.6 Experimental Procedure

At the start of each experiment, participant data (i.e. age, gender, year of university, color vision impairments) was recorded and participants were instructed about the experimental procedure. Participants' answer regarding color vision impairments determined which color scale was used (see Fig. 4). Afterwards, each participant performed 18 needle insertion tasks. The order of tested prototypes was

partially randomized: Prototypes were grouped by the concept factor levels, resulting in three clusters of six prototypes each. The order of clusters was randomized first. The prototypes were then ordered randomly within their cluster.

Before a new concept was introduced, a training phase was conducted where participants could train the insertion task with a randomly selected prototype out of the current cluster and were instructed about the accuracy indicators and accuracy-to-color mapping scales. When participants felt confident with the concept, the six cluster trials began one after another. To begin each trial, participants had to press the aforementioned switch positioned between them and the phantom. Because we did not measure positioning accuracies, participants could freely select a new insertion point on the phantom surface. Projected navigation concepts were moved to the selected insertion points accordingly. Participants then performed the insertion task. Target insertion angle and depth were generated randomly before each trial. Insertion depths could range between 70mm and 90mm. Insertion angles were selected between 0° and 30° at a randomized direction around the axis perpendicular to the insertion site. After each insertion, participants were asked for their subjective ratings on perceived task difficulty regarding the correct adjustment of angle and depth.

## 5 RESULTS

Three-way ANOVAs were conducted for all dependent variables to investigate the three factors' effects. All statistically significant effects are summarized in Table 1. Statistical parameters for non-significant effects are reported in Supplementary Table 1. A full overview of the data is provided in Supplementary Table 2.

### 5.1 Data Exclusion

One participant's data were excluded due to language comprehension issues and resulting performance issues. A technical error occurred with discrete accuracy-to-color mapping: Over-insertion was not color coded in any trials with these implementations. The error was not noticed until after the end of the study. All trials with discrete accuracy-to-color mapping were excluded from the analysis for depth accuracy, task completion time, and subjectively perceived difficulty of depth setting.

### 5.2 Interpretation of Results

This section discusses and attempts to interpret the statistical results that were found for each visualization factor.

#### 5.2.1 Visualization Concept

Results for the Concept factor are summarized in Fig. 7. Main effects across all dependent variables indicate a consistent ranking between the visualization concepts. The Crosshairs concept performed best in all variables. The Circle concept also performed better than the Arrow concept in all dimensions but in absolute depth accuracy. For this indicator, the Circle concept performed significantly worse than both other concepts. The two-way interaction between the Concept factor and the Scaling factor may offer a potential explanation for this. Fig. 6 shows that

TABLE 1
Summary of the ANOVA results (statistically significant effects only)

| Variable / Effect type | Factor (Hypothesis) | df | F | p | $\eta^2$ | Effect | Figure |
|---|---|---|---|---|---|---|---|
| **Angular deviation** | | | | | | | |
| Main effect | *Concept* (H1.1) | 2 | 6.78 | 0.001 | 0.031 | Small effect | Figure 7a |
| **Absolute depth deviation** | | | | | | | |
| Main effects | *Concept* (H1.2) | 2 | 11.37 | <0.001 | 0.053 | Small effect | Figure 7b |
| | *Scaling* (H3.2) | 1 | 106.75 | <0.001 | 0.248 | Large effect | Figure 9b |
| Two-way interaction effect | *Concept * Scaling* (-) | 2 | 5.19 | 0.006 | 0.024 | Small effect | Figure 6 |
| **Task completion time** | | | | | | | |
| Main effects | *Concept* (H1.3) | 2 | 4.64 | 0.011 | 0.030 | Small effect | Figure 7c |
| | *Scaling* (H3.3) | 1 | 17.56 | <0.001 | 0.058 | Small effect | Figure 9c |
| **Subjective difficulty angle** | | | | | | | |
| | *Concept* (H1.4) | 2 | 8.38 | <0.001 | 0.037 | Small effect | Figure 7d |
| Main effects | *Color* (H2.4) | 2 | 3.97 | 0.020 | 0.017 | Small effect | Figure 8c |
| | *Scaling* (H3.4) | 1 | 9.25 | 0.003 | 0.020 | Small effect | Figure 9d |
| **Subjective difficulty depth** | | | | | | | |
| Main effects | *Concept* (H1.5) | 2 | 3.28 | 0.039 | 0.021 | Small effect | Figure 7e |
| | *Scaling* (H3.5) | 1 | 24.20 | <0.001 | 0.078 | Medium effect | Figure 9e |

the Circle concept implementations with linear indicator scaling performed worst. Reviewing the raw data reveals that this poor performance is particularly eminent for Mono color implementations. This may be because, in the Circle concept, depth information converges at the injection site which may be occluded or subject to projection shadows. Furthermore, the information bearing area shrinks as the needle approaches the target depth, i.e. the more accurately the needle is placed, the smaller the area in which information is communicated. This is true for all implementations of the Circle concept and may have contributed to the concept's bad performance regarding depth accuracy.

The Arrow concept performed worst in all variables but the absolute depth deviation. Compared to the other two concepts, the Arrow concept has the disadvantage that angular information and depth information are located separately. The user has to shift visual attention between two locations to receive all relevant information. This may have led to the concept's poor performance across variables.

### 5.2.2 Accuracy-to-Color Mapping

Results for the Color factor are summarized in Fig. 8. No significant effects on objective performance indicators (i.e. accuracy or task completion time) were found. Yet, participants perceived the correct angular orientation of the needle to be easier with color information than without and easier for continuous color mapping than for discrete color mapping. Further research is required for this factor as a considerable part of data had to be excluded from the analysis.

### 5.2.3 Indicator Scaling

Results for the Scaling factor are summarized in Fig. 9. Indicator scaling had a strong impact on insertion depth accuracy. Implementations with logarithmic indicator scaling performed better than those with linear indicator scaling. This may be attributed to the fact that a higher resolution of information is provided as the needle approaches the target insertion depth. However, no equivalent effect was found for angular accuracy.

The interaction effect (Fig. 6) suggests that linear scaling in particular performed better in the Crosshairs concepts

than in the other concepts with regards to depth accuracy. One potential reason for this effect may be that, while the radius of the colored-in area increases linearly, the surface area itself grows more than linearly in these concept implementations. If some participants perceived the surface area of the circle as the accuracy indicator, this may have caused the results of the two scaling implementations to be closer in the Crosshairs concept than in the other concepts. However, this interpretation of the interaction effect is somewhat speculative and requires more research.

Task completion time was longer and participants perceived needle orientation and insertion depth setting to be more difficult using logarithmic scaling implementations than linear scaling implementations. This may also be attributed to the higher resolution of information at small inaccuracies. People may perceive their angular or depth error to be larger than it is. They may then spend more effort (and time) to correct their perceived mistake. They may also experience some frustration over not achieving the same level of accuracy feedback as they achieved with the linear scaling implementations. The increased effort and potential frustration may have led to the higher perceived difficulty.
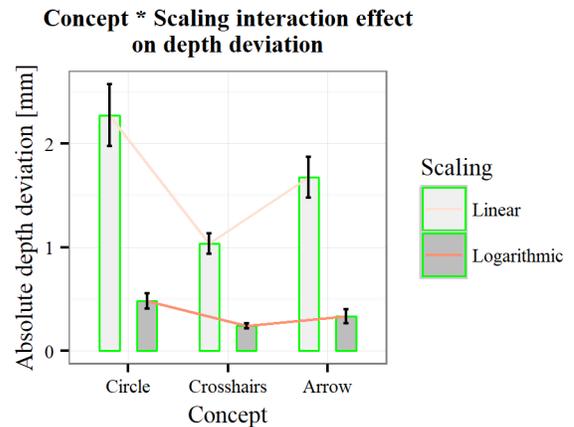


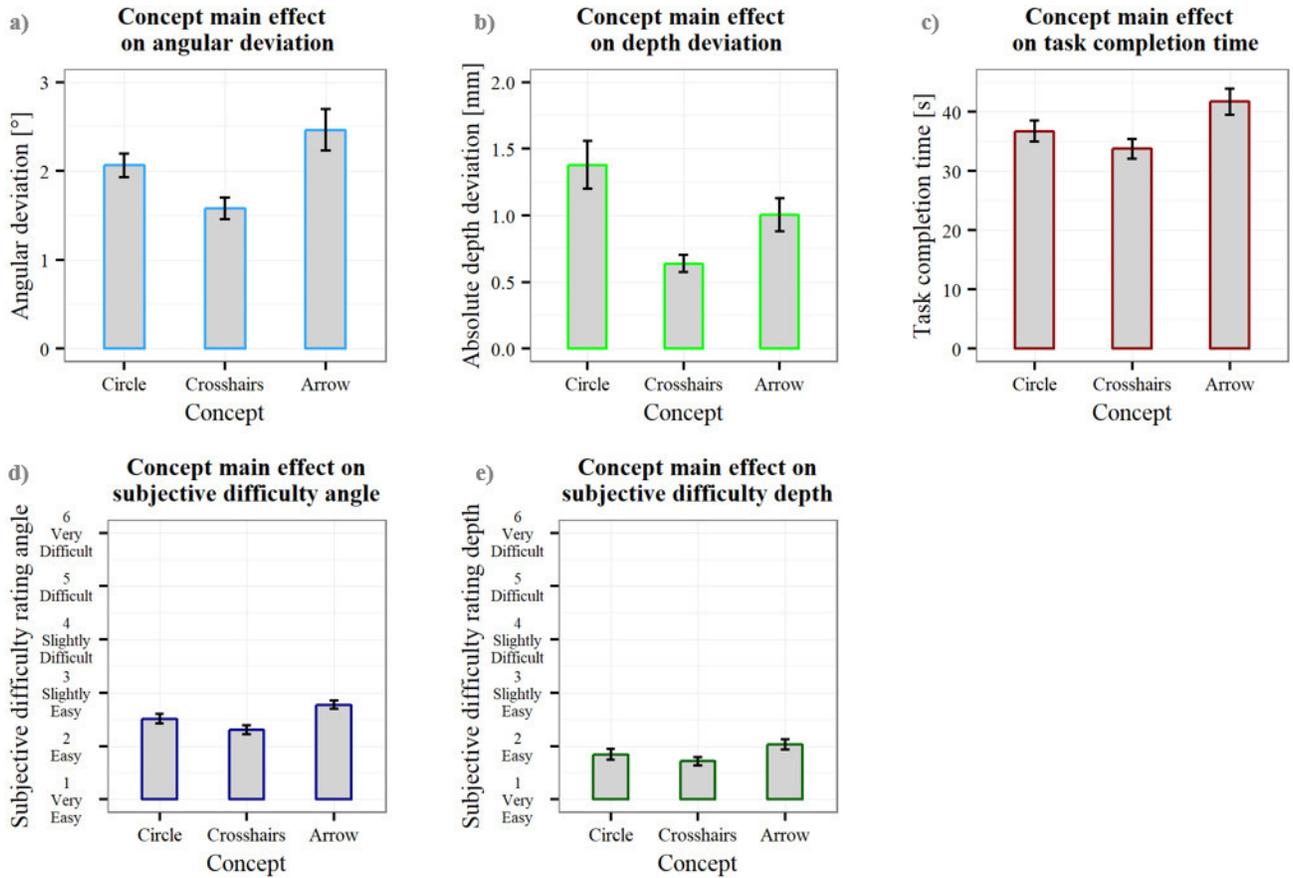Fig. 6. Two-way interaction effect on absolute depth deviation. (Error bars represent standard error.)

Fig. 7. Main effects of the *Concept* factor on: a) angular deviation, b) absolute depth deviation, c) task completion time, d) subjective difficulty angle, and e) subjective difficulty depth. (Error bars represent standard error.)
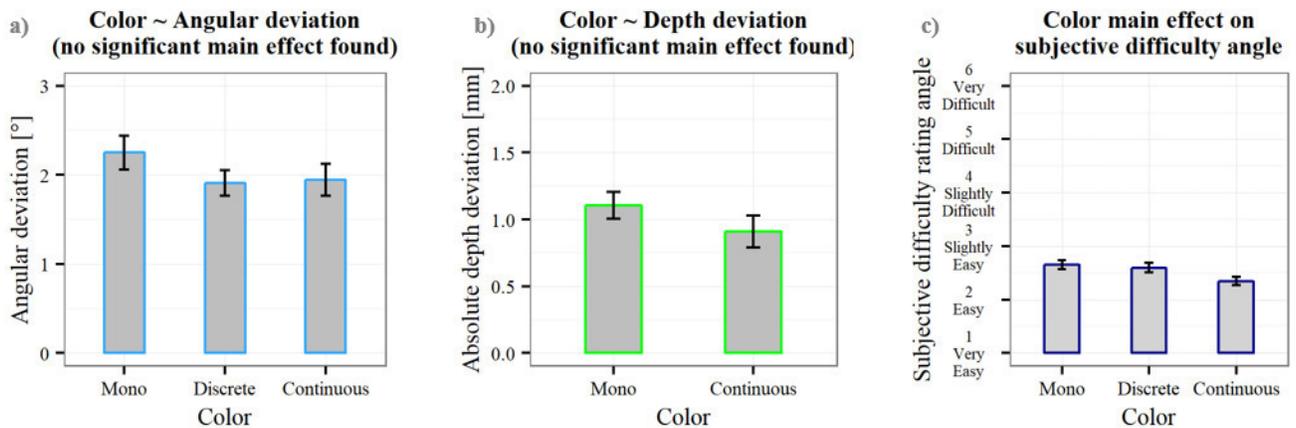


Fig. 8. Main effects of the *Color* factor on: a) angular deviation, b) absolute depth deviation, and c) subjective difficulty angle. (Error bars represent standard error.)
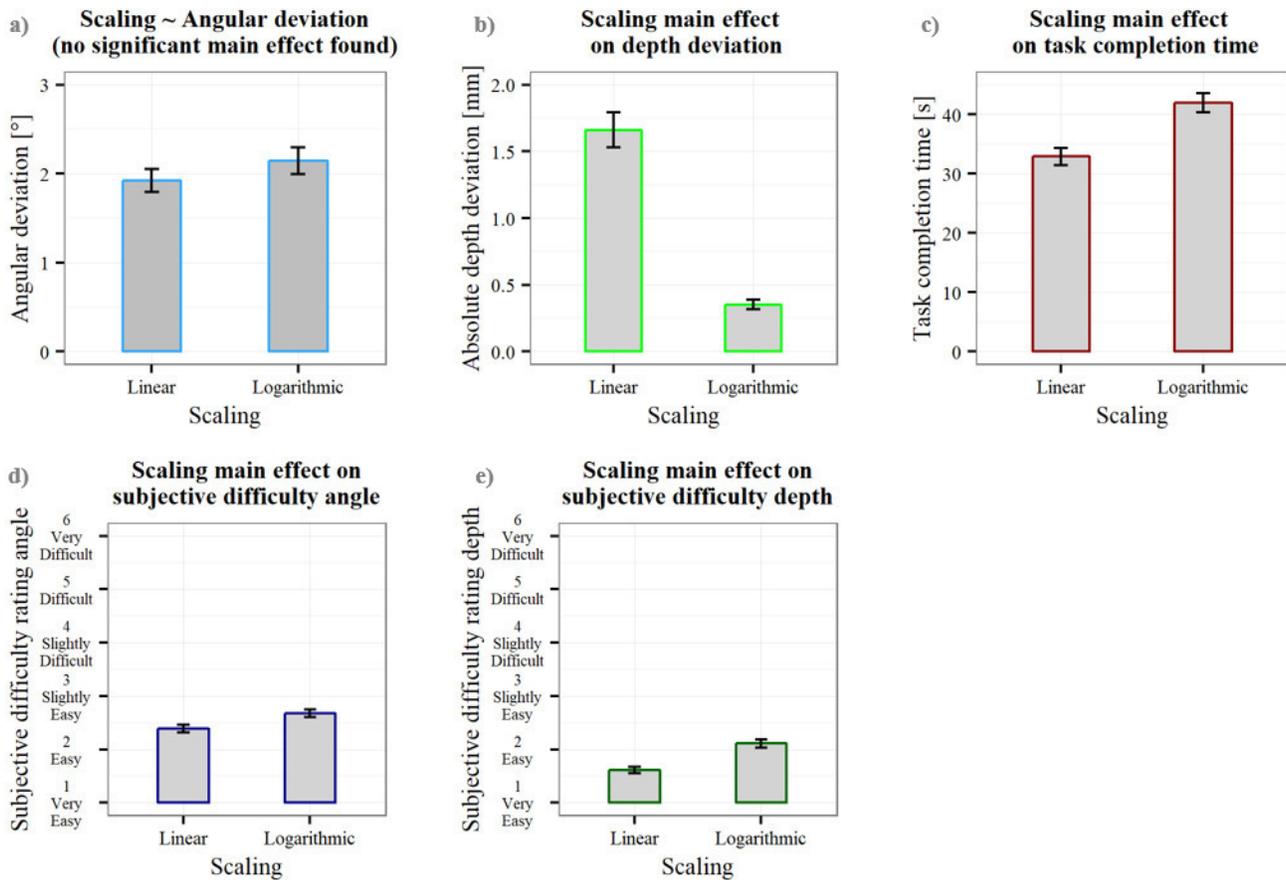
Fig. 9. Main effects of the *Scaling* factor on: a) angular deviation, b) absolute depth deviation, c) task completion time, d) subjective difficulty angle, and e) subjective difficulty depth. (Error bars represent standard error.)

## 6 DISCUSSION

Clear results point towards the conclusion that the Crosshairs concept is superior for both the orientation and the insertion-to-target-depth aspects of the task. While this is consistent across objective and subjective performance indicators, it only allows for drawing conclusions about the overall concepts. We cannot rule out the possibility that the concepts' single dimensions (i.e. angle *or* depth visualization) might have ranked differently. However, this study investigated concepts for a medical application in which angular and depth information are required in an integrated form. Thus, we believe that testing the integrated concepts yielded the most meaningful results. For a better understanding of visualization concepts' effects, more research would be beneficial. Besides different navigation indicator combinations, the evaluation of other visualization concepts would be useful, as well. See-through vision and access path visualizations were excluded from this study but may be advantageous approaches that require separate research activities.

The examination of different indicator scaling methods showed the clearest results in our study. Logarithmic indicator scaling proved to be advantageous over a linear scaling method for the depth indicator. However, prototypes using this technique were more time-consuming and perceived as more difficult. For this study, we preset the curve-varying

parameters b and t (see Equation 1). Further research regarding the effect of different parameters may be beneficial to find an indicator scaling method that is a compromise between accuracy on the one hand and subjective difficulty / task completion time on the other hand.

Regarding accuracy-to-color mapping, no clear conclusions can be made. Because a considerably large amount of data had to be excluded from the analysis regarding this variable, a conclusive comparing study investigating the differences of discrete and continuous color schemes for needle insertion tasks is yet to be conducted. Besides, we could not show that using color is advantageous over using monochromatic visualizations. This may be because changing the size of navigation indicators alone conveyed enough information to correctly align and insert the needle. Thus, indicating angular or depth accuracy via color coding may not have provided enough supplementary information to result in significant differences. The color schemes and accuracy thresholds used for color mapping in this study may have contributed to this.

One disadvantage of using projective AR is the occlusion of information by shadows. This problem may have influenced the use of our navigation prototypes when the participants' hands or the tracked needle instrument itself shadowed the projection site, partly occluding the displayed navigation visualizations. Shadows may have influenced

the use of each concept differently. However, we did not evaluate this potential issue. Further research regarding the impact and implications caused by shadows may be an interesting research objective of future experiments. Other display modalities are not affected by this problem. We chose a projection approach because of promising prior research results. However, recent advances in different modalities might prove to be advantageous. For example, optical see-through HMD approaches have recently been investigated for needle insertion tasks, as well [11], [14]. Research regarding the appropriate choice of display modalities therefore seems to be of great interest. Furthermore, solely focusing visual feedback might not be the ideal approach. Some state of the art approaches suggest fusing visual feedback with auditory support [15]. More research on this topic may help to further improve navigation support.

We only investigated the needle placement accuracy dimensions orientation and insertion depth. Positioning accuracy at the entry point was not regarded but might be a topic worthy of further exploration. The entry point defines the angle and depth required for a precise needle insertion to a specific target and may determine if a safe insertion path is feasible at all. It may be possible to compensate for positioning deviations with angular and depth adjustments, but risk structures might occlude a clear access path. Mewes et al. compared two different entry point visualizations but could not show statistically significant differences [24]. Further research on this topic may improve the understanding of entry point visualization concepts.

We neglected the potential effects of moving target positions due to respiration. Some procedures are conducted under respiratory arrest or on organs only marginally affected by respiratory movement (e.g. the spine or skull). However, organ shift due to respiration is a major issue for various procedures and thus may limit our results applicability if it does not get considered. Moreover, target and needle deformation are additional factors influencing final positioning accuracy. To overcome these issues, various methods and models for respiratory movement [27] as well as organ and instrument deformation estimation [28] have been developed. Such information can be used to adjust displayed navigation information. Additionally, we focused on free-hand insertion tasks only. Other needle-based interventions are carried out semi-automatically or fully automatically using robotics. This paper's evaluated methods may contribute to these procedures, as well. If a human operator is in control of the robot, navigation support is still required. Additionally, projective AR may contribute valuable feedback of the current insertion state to observers [29].

Generally, the relative, comparative results in this study were the main focus of research. Absolute accuracy values convey very little information regarding clinical aptitude of the tested concepts because this is determined by absolute accuracy of the final needle position. This, in turn, is determined by three factors: Insertion point placement, needle orientation and insertion depth. Moreover, we did not focus on system calibration for high technical (i.e. tracking and projection) accuracy which may add bias regarding absolute needle position accuracy. Thus, clinical aptitude of an integrated prototype should be evaluated in a separate study.

# 7 CONCLUSION

This work investigated the performance of various projective AR navigation visualization concepts and implementations in a simulated medical needle insertion task. The implementations differed in visualization concept, navigation indicator scaling, and accuracy-to-color mapping approaches. Performance was determined by objective parameters (accuracy and task completion time) and users' subjective perception of task difficulty.

The evaluation study yielded meaningful results regarding the visualization concepts and indicator scaling which can be translated into clear design recommendations. Thus, our results contribute to the future design of projective AR support in needle navigation tasks. We furthermore identified a range of open research questions which we believe should be addressed to further advance the field of projective AR support in medical and potentially a broader field of practical applications.

We believe that a meaningful follow-up of this work will be the implementation of an integrated prototype, following the insights from this study and including technical and needle placement accuracy. This prototype could then be tested to evaluate if overall needle insertion accuracy with the resulting projective AR navigation support can achieve better results than current solutions.

## REFERENCES

[1] 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," *Ann Surg*, vol. 248, no. 3, pp. 475–486, Sep 2008.

[2] 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," *Comput Aided Surg*, vol. 16, no. 4, pp. 181–187, 2011.

[3] F. Fischbach, J. Bunke, M. Thormann, G. Gaffke, K. Jungnickel, J. Smink, and J. Ricke, "MR-guided freehand biopsy of liver lesions with fast continuous imaging using a 1.0-T open MRI scanner: experience in 50 patients," *Cardiovasc Intervent Radiol*, vol. 34, no. 1, pp. 188–192, Feb 2011.

[4] J. Engstrand, G. Toporek, P. Harbut, E. Jonas, H. Nilsson, and J. Freedman, "Stereotactic CT-Guided Percutaneous Microwave Ablation of Liver Tumors With the Use of High-Frequency Jet Ventilation: An Accuracy and Procedural Safety Study," *AJR Am J Roentgenol*, vol. 208, no. 1, pp. 193–200, Jan 2017.

[5] R. F. Grasso, E. Faiella, G. Luppi, E. Schena, F. Giurazza, R. Del Vescovo, F. DAgostino, R. L. Cazzato, and B. B. Zobel, "Percutaneous lung biopsy: comparison between an augmented reality ct navigation system and standard ct-guided technique," *Int J Comput Assist Radiol Surg*, vol. 8, no. 5, pp. 837–848, 2013.

[6] 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," *Comput Aided Surg*, vol. 16, no. 5, pp. 203–219, 2011.

[7] 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," *Int J Med Robot*, vol. 5, no. 3, pp. 297–308, 2009.

[8] C. Hansen, D. Black, C. Lange, F. Rieber, W. Lamadé, M. Donati, K. J. Oldhafer, and H. K. Hahn, "Auditory support for resection guidance in navigated liver surgery," *Int J Med Robot*, vol. 9, no. 1, pp. 36–43, 2013.

[9] 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," *Int J Comput Assist Radiol Surg*, vol. 7, no. 4, pp. 547–556, 2012.

[10] 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," *Int J Comput Assist Radiol Surg*, vol. 11, no. 1, pp. 107–117, 2016.

[11] 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*, vol. 3, no. 4, pp. 4156–4162, 2018.

[12] 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*, vol. 240, no. 1, pp. 230–235, Jul 2006.

[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," *AJR Am J Roentgenol*, vol. 198, no. 3, pp. W266–273, Mar 2012.

[14] P. K. Kanithi, J. Chatterjee, and D. Sheet, "Immersive augmented reality system for assisting needle positioning during ultrasound guided intervention," in *Proceedings of the 10th Indian Conference on Computer Vision, Graphics and Image Processing*, 2016, p. 65.

[15] 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 *2015 IEEE International Symposium on Mixed and Augmented Reality*, 2015, pp. 7–12.

[16] T. Kuzhagaliyev, N. T. Clancy, M. Janatka, K. Tchaka, F. Vasconcelos, M. J. Clarkson, K. Gurusamy, D. J. Hawkes, B. Davidson, and D. Stoyanov, "Augmented reality needle ablation guidance tool for irreversible electroporation in the pancreas," in *Medical Imaging*, vol. 10576, 2018, p. 1057613.

[17] R. Khlebnikov, B. Kainz, J. Muehl, and D. Schmalstieg, "Crepuscular rays for tumor accessibility planning," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2163–2172, 2011.

[18] J. Alpers, C. Hansen, K. Ringe, and C. Rieder, "Ct-based navigation guidance for liver tumor ablation," in *Eurographics Workshop on Visual Computing for Biology and Medicine.* The Eurographics Association, 2017.

[19] 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*, vol. 238, no. 2, pp. 497–504, 2006.

[20] 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," *Int J Radiat Oncol Biol Phys*, vol. 70, no. 3, pp. 944–952, 2008.

[21] J. Kreiser, J. Freedman, and T. Ropinski, "Visually supporting multiple needle placement in irreversible electroporation interventions," in *Computer Graphics Forum*, vol. 37, 2018, pp. 59–71.

[22] A. Seitel, L. Maier-Hein, S. Schawo, B. Radeleff, S. Mueller, F. Pianka, B. Schmied, I. Wolf, and H.-P. Meinzer, "In-vitro evaluation of different visualization approaches for computer assisted targeting in soft tissue," *Int J Comput Assist Radiol Surg*, vol. 2, pp. S188–S190, 2007.

[23] W. Y. Chan and P. A. Heng, "Visualization of needle access pathway and a five-DoF evaluation," *IEEE J Biomed Health Inform*, vol. 18, no. 2, pp. 643–653, Mar 2014.

[24] A. Mewes, F. Heinrich, B. Hensen, F. Wacker, K. Lawonn, and C. Hansen, "Concepts for augmented reality visualisation to support needle guidance inside the MRI," *Healthc Technol Lett*, vol. 5, no. 5, pp. 172–176, Oct 2018.

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

[26] A. Mewes, F. Heinrich, U. Kägebein, B. Hensen, F. Wacker, and C. Hansen, "Projector-based augmented reality system for interventional visualization inside mri scanners," *Int J Med Robot*, p. e1950, 2018.

[27] L. Maier-Hein, S. A. Müller, F. Pianka, A. Seitel, B. P. Müller-Stich, C. N. Gutt, U. Rietdorf, G. Richter, H.-P. Meinzer, B. M. Schmied, and I. Wolf, "In-vitro evaluation of a novel needle-based soft tissue navigation system with a respiratory liver motion simulator," in *Medical Imaging*, vol. 6509, 2007, p. 650916.

[28] P. Moreira and S. Misra, "Biomechanics-Based Curvature Estimation for Ultrasound-guided Flexible Needle Steering in Biological Tissues," *Ann Biomed Eng*, vol. 43, no. 8, pp. 1716–1726, Aug 2015.

[29] C. Vogel, M. Poggendorf, C. Walter, and N. Elkmann, "Towards safe physical human-robot collaboration: A projection-based safety system," in *International Conference on Intelligent Robots and Systems*, 2011, pp. 3355–3360.