



# Multimodal markers for technology-independent integration of augmented reality devices and surgical navigation systems

Mohamed Benmahdjoub<sup>1,2</sup> · Wiro J. Niessen<sup>1,3</sup> · Eppo B. Wolvius<sup>2</sup> · Theo van Walsum<sup>1</sup>

Received: 13 November 2020 / Accepted: 13 April 2022 / Published online: 27 May 2022  
© The Author(s) 2022

## Abstract

Augmented reality (AR) permits the visualization of pre-operative data in the surgical field of view of the surgeon. This requires the alignment of the AR device's coordinate system with the used navigation/tracking system. We propose a multimodal marker approach to align an AR device with a tracking system: in our implementation, an electromagnetic tracking system (EMTS). The solution makes use of a calibration method which determines the relationship between a 2D pattern detected by an RGB camera and an electromagnetic sensor of the EMTS. This allowed the projection of a 3D skull model on its physical counterpart. This projection was evaluated using a monocular camera and an optical see-through device (HoloLens 2) (<https://www.microsoft.com/en-us/hololens/>) achieving an accuracy of less than 2.5 mm in the image plane of the HoloLens 2 (HL2). Additionally, 10 volunteers participated in a user study consisting of an alignment task of a pointer with 25 projections viewed through the HL2. The participants achieved a mean error of 2.7 1.3 mm and 2.9 2.9° in positional and orientation error. This study showcases the feasibility of the approach, provides an evaluation of the alignment, and finally, discusses its advantages and limitations.

**Keywords** Augmented reality · Mixed reality · Navigation system · Calibration · Alignment · Surgical interventions · Feasibility studies

## 1 Introduction

In the last decade, the usage of surgical navigation systems for cranial (Eggers et al. 2009), neurosurgical (Hara et al. 2020; Li et al. 2018) or liver surgery (Banz et al. 2016) has grown considerably. Navigation systems guide the surgeon

intraoperatively by aligning patient-specific data (CT-scan, MRI-scan) to the patient, and by visualizing relevant information on the screen. The use of accurate tracking systems and software tools that control and visualize patient-specific models and plannings, improved surgical outcomes when compared to conventional freehand techniques (Chotana-phuti et al. 2008).

Based on the tracking technology used, surgical navigation systems can be divided in two groups: those based on optical tracking systems (OTS) and those based on electromagnetic tracking systems (EMTS) (Cleary et al. 2010). For systems using optical tracking, cameras (RGB or infrared) capture landmarks, such as natural feature patterns or reflective markers, in their field of view. These landmarks are arranged in a known geometry allowing the system to compute the 6 degrees-of-freedom (6 DoF) pose of the sensor, and subsequently, of the object to which the sensor is attached. For systems using electromagnetic tracking, a field generator generates an electromagnetic field within a specific spatial volume. Sensors (coils) that are positioned inside this space can be localized accurately.

✉ Mohamed Benmahdjoub  
m.benmahdjoub@erasmusmc.nl

Wiro J. Niessen  
w.niessen@erasmusmc.nl

Eppo B. Wolvius  
e.wolvius@erasmusmc.nl

Theo van Walsum  
t.vanwalsum@erasmusmc.nl

<sup>1</sup> Biomedical Imaging Group Rotterdam, Department of Radiology & Nuclear medicine, Erasmus MC, 3015 GE Rotterdam, The Netherlands

<sup>2</sup> Department of Oral and Maxillofacial Surgery, Erasmus MC, 3015 GE Rotterdam, The Netherlands

<sup>3</sup> Department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

Three limitations have been reported in the literature regarding the use of navigation systems (Mezger et al. 2013). The first limitation is the hand-eye coordination: For instance, it is not uncommon that when a surgeon moves his drill in a vertical direction, the drill on screen appears as moving in a horizontal direction. The second limitation is the need to continuously switch focus between looking at the screen, and the surgical site on the body. The third limitation is the depth perception: Visualizing 3D anatomical structures on a 2D screen requires the operator to mentally reconstruct the 3D shapes based on flat 2D images. All of these issues lead to cognitive overload which makes the use of navigation systems difficult for surgeons, especially those in training. Augmented reality (AR) has been used previously to support in various applications, such as assisting with assembly tasks (Westerfield et al. 2015; Khenak et al. 2020), to address the aforementioned limitations. It allows the integration of virtual objects, such as surgical plannings or critical structures, in the field of view of the user. In the medical field, the application of AR intraoperatively is still a research topic. This is due to ergonomics and accuracy requirements in surgical procedures. Furthermore, even though substantial progress has been made in the field of AR in the past decade, some challenges still need to be addressed. The most important one is the registration or alignment challenge. It consists of accurately superimposing a 3D model on its corresponding physical one. In the past, two types of systems have been developed for medical AR: standalone AR systems and navigation-systems-based AR systems (Benmahdjoub et al. 2020). In the first category, the AR device (e.g., headset) is responsible for acquiring the scene, through which it computes the 6 DoF pose of the objects to augment, subsequently providing an adequate visualization. For pose computation and registration, three approaches have been described: marker based (Badiali et al. 2014; Zhu et al. 2018; Gsaxner Christina et al. 2019), where recognizable fiducials are attached to the physical phantom; markerless (Wang et al. 2014), where feature extraction from the image is required to determine the pose; or manual in some cases where the 3D model is manually placed in its intended position (Incekara et al. 2018).

In the second category, a combination of navigation systems and AR devices is used. Here, the navigation system is responsible for tracking the patient and the instruments involved in the operating room (OR), whereas the AR device is used for projecting and visualizing the required anatomical structures (Zinser et al. 2013; Gavaghan et al. 2012; Meulstee et al. 2019; Kuzhagaliyev et al. 2018; Chen et al. 2015). In this category, the integration of AR within a navigation system requires the AR device's coordinate system  $D$  to be aligned with the navigation system's coordinate systems  $N$ . The common way to achieve this alignment is by attaching the navigation system's markers/sensors to the AR device.

This way, the navigation system can compute the position and orientation of the AR device. Afterward, by performing a calibration procedure,  $D$  can be aligned on  $N$  (Gavaghan et al. 2012; Chen et al. 2015; Meulstee et al. 2019). This approach does not work if the navigation system is based on electromagnetic (EM) tracking. In EMTS, sensors are generally wired, and the tracking volume of commonly used systems is too small (around  $50\text{cm}^3$ ). Furthermore, AR devices, instruments and the OR equipment may distort the EM field.

Compared to both camera and smartphone, optical see-through (OST) devices require an additional calibration regardless of the method of AR integration. Eye calibration is crucial for AR applications requiring model alignment. It is defined as the procedure that allows to compute the transformation between the device's tracking system (e.g., RGB camera) and the user's eyes. Multiple methods have been presented in the literature (Itoh et al. 2014; Genc et al. 2002; Makibuchi et al. 2013; Azimi et al. 2017). This procedure in general requires the user to align virtual elements, seen through the headset, with real locations in the world's coordinate system, to estimate the required transformation (Tuceryan et al. 2000). Because of this fundamental difference between OST-based AR and video-based AR, the evaluation of the projected models' placements in the real world is different; since only users perceive the final output, traditional error measurements on 2D images are not enough. Generally, a user study is conducted to assess OST-based AR systems (Azimi et al. 2020).

Currently, the combination of AR and conventional navigation systems has predominantly been implemented for optical tracking systems. The state-of-the-art approach requires adapting the AR device by calibrating it with the tracking system after rigidly attaching sensors to it. The purpose of this study is to develop and assess a generic approach which aligns an AR device, such as the HoloLens 2 (HL2), with existing navigation systems. This approach relies on an assembled multimodal marker which consists of a 2D pattern and a sensor attached to it. The alignment method's feasibility is demonstrated using an EMTS and a smartphone and then evaluated using a monocular (RGB) camera, and an OST device HL2.

The approach is explained, technically evaluated, and discussed.

## 2 Methods and materials

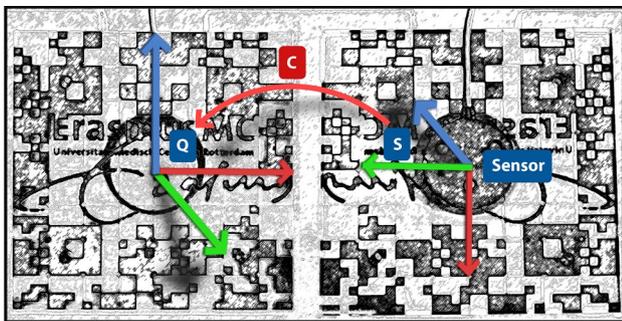
The proposed solution integrates an AR device in a navigation system without being tracked by the latter. Rather, instead of tracking the AR device with the navigation system, a reference object is introduced trackable by both the camera of the AR device and the navigation system. The different sensor modalities (camera of the AR device and

an OTS/EMTS) imply that the features tracked by the two modalities differ and thus require two different types of markers which leads to our multimodal marker solution (Fig. 1). In the following section, the concept of the multimodal marker is presented which is independent of the underlying tracking technology of the navigation system.

### 2.1 Multimodal marker

The multimodal marker is composed of two markers: a marker with features trackable by the navigation system (reflective spheres, coils,...) through their respective tracking device, and a marker containing features trackable by the AR-device (2D pattern) through its RGB camera. Figure 1 presents an example of such a multimodal marker. The AR marker (2D pattern) defines a local reference coordinate system  $Q$ , which can be detected by the AR-device, and thus also implicitly defines the AR device’s coordinate system. The sensor of the navigation system defines a local coordinate system  $S$ , which is related to the navigation system’s coordinate system  $N$ . If the spatial relationship between  $Q$  and  $S$  is known, the coordinate system of the AR device can be linked to the coordinate system of the navigation system. This allows the display of the 3D models from the navigation system in the AR device.

Therefore, a calibration procedure needs to be performed to determine the rigid transformation  $C$  between the two coordinate systems (Fig. 1). In the following sections, we will describe this calibration procedure, and also how it facilitates augmentation of the instruments and the patient.



**Fig. 1** Multimodal marker—Left: 2D printed pattern with its local coordinate system  $Q$ . Right: The other side of the multimodal marker: a circular shaped electromagnetic sensor with its local coordinate system  $S$  attached rigidly to the 2D pattern.  $C$  is the calibration matrix which aligns  $S$  on  $Q$

### 2.2 Calibration

The calibration process determines the rigid transformation  $C$  from the local coordinate system of the 2D pattern  $Q$  to the local coordinate system of the sensor  $S$  defined in its field generator’s coordinate system  $N$  (Fig. 1). The steps for this procedure are as follows (see also 1 for the annotations):

1. Define by construction  $n$  virtual points  $L_V = \{v_1, v_2, \dots, v_n\}$  in  $Q$  (Fig. 1).
2. Define the equivalent physical points  $L_P = \{p_1, p_2, \dots, p_n\}$  which can be acquired using a pointer localized in  $N$  by the tracking system of the navigation system (Fig. 1). Each point  $p_i$  represents the position given by a  $4 \times 4$  transformation matrix of the pointer  $T_O^N$  in the tracking system’s coordinate system.
3. Compute  $L_S = \{s_1, s_2, \dots, s_n\}$  obtained by transforming  $L_P$  to the sensor’s local coordinate system  $S$ :

$$p_{i,S} = T_S^{N-1} p_i \tag{1}$$

where  $p_{i,S}$  is  $p_i$  defined in  $S$  and  $T_S^N$  is the sensor’s pose in  $N$ .

4. Use a point-based registration method to compute the landmark rigid transformation  $C$  which transforms  $L_S$  to  $L_V$  (Horn and Berthold 1987).

### 2.3 Instruments augmentation

Navigation with AR is possible when the multimodal marker is added to the scene and puts in the proximity of the target object to augment. The multimodal marker needs to be visible by both the tracking system (sensor inside the tracking volume) and the AR device (pattern) simultaneously.

In this scenario, the following equation represents the set of transformations to align any sensor defined in  $N$  into the local coordinate system of the marker  $Q$ :

$$T_I^Q = C_S^Q T_S^{N-1} T_I^N, \tag{2}$$

where  $T_I^N$  is the instrument’s pose (i.e., pointer) in  $N$ ,  $T_S^N$  is the sensor’s pose in  $N$ , and  $T_I^Q$  is the final transformation attributed to the virtual model in the local coordinate system  $Q$  of the 2D pattern.

If the AR device uses a monocular RGB camera, the necessary projection needs to be applied to go from the 3D space to the 2D image by applying the following equation:

$$T_I^{RGB} = K_{RGB} A_Q^{Camera} T_I^Q, \tag{3}$$

where  $A_Q^{Camera}$  is the camera pose estimation based on the 2D pattern, and  $K_{RGB}$  is the intrinsic parameters matrix of the

camera (which can be obtained through a calibration procedure).

## 2.4 Patient augmentation

The augmentation of the patient by a 3D model of the anatomical structures requires an initial image-to-patient registration to align the patient model with the patient, and the subsequent update of this alignment according to the patient's movements. To achieve the latter, a sensor is attached to the patient after which an image-to-patient registration is performed. To this end, the registration points are determined in the patient sensor's coordinate system  $\mathbf{P}$  as follows:

$$p_{i,P} = T_P^{N-1} t_i \quad , \quad (4)$$

where  $T_P^N$  is the patient sensor's pose in  $N$ ,  $t_i$  the registration point  $i$  in  $N$ , and  $p_{i,P}$  the physical registration point  $i$  defined in  $\mathbf{P}$ . Subsequently, the transformation  $\mathbf{R}$  between the image registration points and the ones defined in  $\mathbf{P}$  is computed.

The patient augmentation positions the patient 3D model at the right location with respect to the 2D pattern's coordinate system  $\mathbf{Q}$ :

$$R_I^Q = C_S^Q T_S^{N-1} T_P^N R_I^P \quad , \quad (5)$$

where  $R_I^P$  is the registration matrix from the patient model to  $\mathbf{P}$ ,  $T_P^N$  is the pose estimation of the sensor which is attached to the patient in  $N$ ,  $T_S^N$  is the pose estimation of the multimodal marker's sensor in  $N$ , and  $C_S^Q$  is the computed calibration matrix between the multimodal marker's sensor and its pattern.  $R_I^P$  can then replace  $T_I^Q$  in Eq. 3 for the projection of the image-based model.

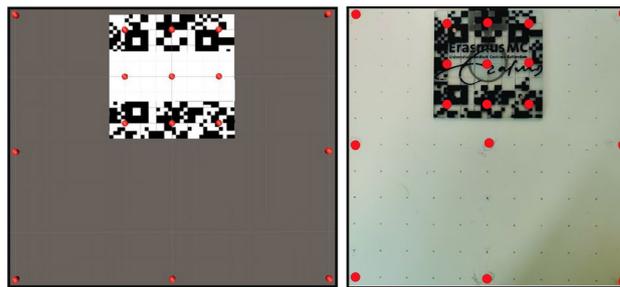
The complete set of transformations is illustrated in Fig. 3.

## 3 Experiments

### 3.1 System overview and implementation

The multimodal marker setup was implemented in C# and Unity (including the communication protocol between the PC and the tracking system) and can be found in [<https://gitlab.com/radiology/igit/ar/ar-em>]. In the experiments, we used:

1. An Aurora V2 electromagnetic navigation system (Northern Digital Inc., Canada) comprising of a field generator, a control unit, one electromagnetic pointer and two coils (sensors)  $\mathbf{S}$  and  $\mathbf{P}$ .

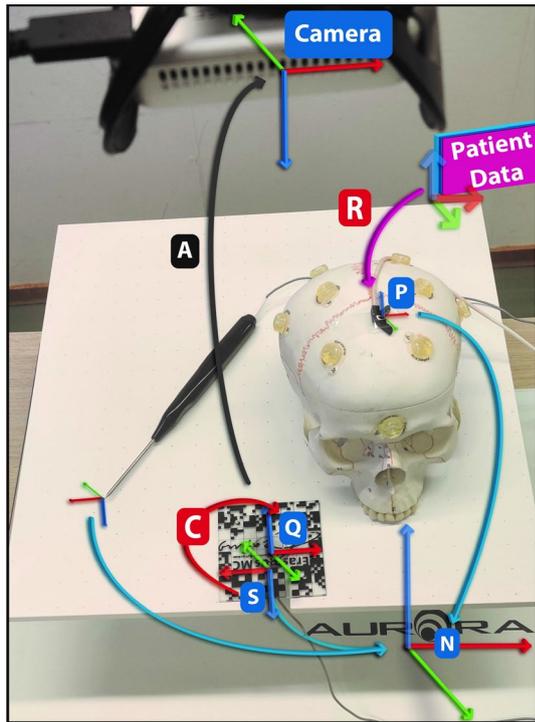


**Fig. 2** In red, an example of possible virtual calibration points (left) and their physical correspondences (right)

2. A 2D pattern engraved on a plate of  $80 \times 80$  mm (Fig. 1). In addition, 49 divot points were drilled on it (10 mm distance between adjacent points) to be used as points for calibration and assessment. All the divots have known coordinates with respect to the local coordinate system  $\mathbf{Q}$ .
3. A calibration board of  $400 \times 400$  mm size with divots each 20 mm (Fig. 3). It is used for calibration and/or accuracy measurements. The divot drilling on both the calibration board and the 2D pattern was performed using the submillimeter numerically controlled drilling machine Fehlmann Picomax 60M (Weissach-Flacht, Switzerland).
4. A skull phantom (3B Scientific, Hamburg, Germany) with pinpoint markers (PinPoint #128, Beekley Medical, Bristol, USA).
5. An RGB camera Logitech Brio (Lausanne, Switzerland) with a resolution of 1080p and a field of view of  $65^\circ$  to capture the scene.
6. Microsoft HoloLens 2.
7. (Vuforia 2020) as a pose estimation tool and a 2D pattern tracking module.
8. (MevisLab 2020) for manual annotation of the registration points on the CT data, and the extraction of the 3D model (.obj) file of the skull.

A sensor  $\mathbf{S}$  is attached to the 2D pattern plate, and the resulting multimodal marker is rigidly attached to the board, rigidly for reproducible calibration, whereas another sensor  $\mathbf{P}$  is attached to the 3D printed model to track the position of the skull. Figure 3 illustrates the experimental setup.

The setup's architecture while running is presented in Fig. 2. To demonstrate this architecture (Sect. 3.6), the following components were used: a desktop computer connected to an EMTS and a monocular RGB camera, functioning as an AR device and as a server for the secondary device; a smartphone Oneplus 7 pro (Shenzhen, China)/Microsoft HL2 as a secondary AR device; a serial communication



**Fig. 3** Overview of the setup: The field generator (*N*), a white board on top of it, the black pointer, a 3D printed skull with a sensor attached on top (*P*), a multimodal marker and a Camera (*Q* and *S*). The arrows represent the transformations between the coordinate systems (Table 1)

**Table 1** Annotations used in the manuscript

Symbols	Meaning
$T_B^A$	Transformation from B to A
<i>O</i>	CS of the EM pointer
<i>P</i>	CS of the EM sensor attached to the patient
<i>Q</i>	CS of the multimodal marker
<i>S</i>	CS of the ME sensor of the multimodal marker
<i>N</i>	CS of the field generator
<i>C</i>	Calibration transformation from <i>S</i> to <i>Q</i> : $T_S^Q$
<i>A</i>	Pose estimated by the RGB camera
<i>R</i>	Image to patient registration transformation: $T_I^P$

protocol (from the navigation system to the desktop); and a TCP/IP connection from the desktop to the secondary AR device.

### 3.2 Pointer pose

To assess the pointer pose, the board on which the divots were drilled was used. The distances between the divots are known and used as ground truth for the measurement. Pairs

of divots (100 mm apart) were touched with the pointer tip, and their intra-distances were compared with the reference standard. This was done for twelve pairs of points (18 divots were acquired). The mean error was computed as follows:

$$E_p = \frac{1}{m} \sum_{i=1}^n |gd - d(p_i, p_{i+1})| \quad (6)$$

where  $p_i$  is the estimated position using the pointer,  $p_{i+1}$  is the next adjacent divot estimated by the pointer,  $d$  is the Euclidean distance,  $m$  is the number of measurements (12) and  $gd$  is the known distance, by construction, of the board (100 mm).

### 3.3 Calibration assessment

The accuracy of several calibrations was assessed. For each calibration, the distances between positions determined with the pointer and transformed by the calibration, and the same points determined on the ground truth divots (virtual points on the multimodal marker and the board) were computed. To this end, the marker with the divots was rigidly fixed to the calibration board in a reproducible manner with an accurately known position. As a consequence, the positions of the divots in the board were known with respect to the multimodal marker. For each scenario, the calibration matrix was computed. Next, a pattern of divots on the board was pinpointed and transformed using the calibration matrix to convert these divots to the multimodal marker’s coordinate system *Q* (see Eq. 2). The error was computed as the distance between the transformed measured position, and the reference position in the ground truth divots (defined in *Q* by construction). Measurements were acquired for nine scenarios. For each of the following calibration surfaces 60 × 60 mm, 200 × 200 mm and 400 × 400 mm, a calibration based on 4, 12 and 24 points was performed and assessed. Each of the nine calibration scenarios was performed five times.

### 3.4 Marker tracking and instrument augmentation

The HL2 was used for the overlay assessment on the board based on marker tracking only (no navigation system or calibration included). In addition, based on calibration C3 (Table 3), the pointer accuracy was evaluated (calibration included).

The overlay on the calibration board using marker tracking only was performed as follows:

1. Capturing an image B1 of the evaluation board scene.
2. Capturing a second image B2 of the augmented evaluation board (given the known position of divots w.r.t. the multimodal marker by construction).

3. Annotating the divots  $D$  (used for calibration assessment) on B1.
4. Annotating the directly adjacent divots (to the evaluation divots used)  $Ad$  on B1.
5. Annotate the projected divots  $Pd$  on B1.
6. Computing the shift on each evaluation divot in mm as follows:

$$\beta_i = \frac{gd \times d(D_i, Pd_i)}{d(D_i, Ad_i)} \quad (7)$$

where  $\beta_i$  is the shift in millimeter between the divot  $i$  and its projected representation.  $gd$  is the known, by construction, ground truth distance between a divot and its closest adjacent divot (20mm),  $d$  is the Euclidean distance,  $D_i$  is the divot  $i$ ,  $Pd_i$  is the projected divot  $i$ ,  $Ad_i$  is the adjacent divot  $i$ .

Similarly, the pointer augmentation is assessed by following the previous steps and replacing  $P_d$  in 5 with an annotation of the pointer's projected tip (Fig. 3).

### 3.5 Overlay accuracy assessment

In the next set of experiments, the overlay accuracy was assessed. To this end, the 3D model of the skull (obtained from a CT-scan) was registered using the attached markers and subsequently projected on top of the physical printed skull. The following scenarios were considered:

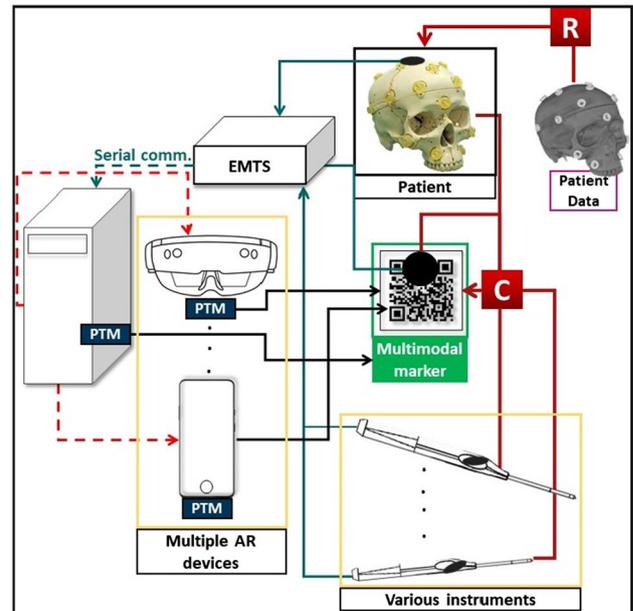
1. Three calibrations (out of 45 calibrations):
  - (a) C1: 60 × 60 mm - 4 calibration points,
  - (b) C2: 400 × 400 mm - 24 calibration points,
  - (c) C3: 400 × 400 mm - 24 calibration points.

These calibrations were chosen based on their RMSE (2.64 mm, 1.60 mm and 1.20 mm, respectively) to investigate the impact of calibration accuracy on the final overlay.

2. Three camera poses with respect to the multimodal marker at 90°, 45° and 20°.

In these experiments, the distance of the camera to the 2D marker was about 50-60cm, which is representative for a clinical navigation setup.

Ten pairs of images were acquired for each camera pose-calibration combination (Fig. 4). Each pair had a different skull pose. The poses were kept the same for the three calibrations. One pair consisted of a non-augmented picture P1, and one augmented picture P2 containing the projected centers of the pinpoint markers (yellow) which were manually annotated on the CT image (Fig. 5). The error metric was



**Fig. 4** The components and their connections in the setup. PTM: pattern tracking module (in our experiments Vuforia). Black circles (ellipses): electromagnetic sensors which are tracked by the EMTS and their pose is sent to the desktop via a serial communication protocol. The complete chain of transformations (Eq. 2), including calibration C and registration R, is computed by the desktop and sent wirelessly (red dashed arrows) to the AR devices. Each AR device is responsible for tracking the pattern, rendering, and projecting the 3D model

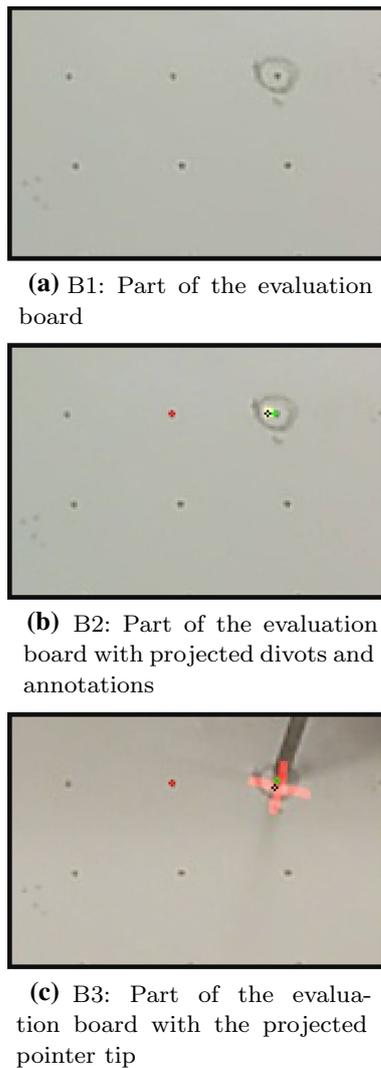
defined as the distance (annotation based) between projected and physical centers of the pinpoint spots using P1 and P2.

For this metric, P2 images were used: knowing the diameter of the registration spots (15 mm), the pixel size in mm in each image was computed (locally for each spot). To locate the target projection points, four registration spots were selected in each image based on the easiness with which they can be annotated, as well as their distribution over the skull (centered).

The registration points were located at the center of each spot's base. Consequently, eight contour points were annotated manually around the base of each torus (spot). An ellipse was fitted to the contour points. The maximum of the width and the height of each ellipse was considered the spot's real diameter in pixels (Fig. 5), whereas the center of the ellipse was considered the target (expected) projection point (Fig. 5). The error  $\epsilon$ , i.e. the distance from the projected point to the ellipse's center was computed as follows:

$$\epsilon = |e_c - p_c| \times s/e_d \quad (8)$$

where  $e_c$  is the ellipse's center,  $p_c$  is the projected center (both in pixels),  $s$  is the registration spot size in mm (15), and  $e_d$  is the fitted ellipse's largest diameter.



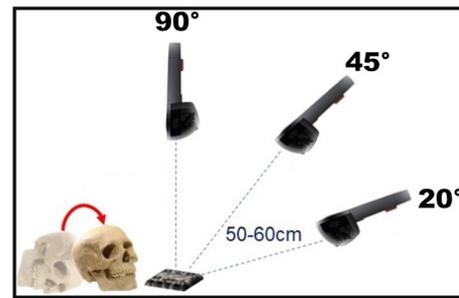
**Fig. 5** Images used for the overlay assessment. Annotations: green circle for the divot to be projected; red circle for the adjacent divot; black circle for the center of the projected divot (projection in yellow), or the projected pointer tip

The mean over each set of ten pictures representing a calibration (C1, C2, or C3) in a given camera pose was computed for both the webcam and the HL2 (3).

An intra-observer evaluation was performed for the ellipse annotation-based assessment to observe the observer-dependency in this metric. Three operators were asked to annotate four registration spots on a set of 30 images.

### 3.6 System output assessment

The system's output (Sect. 3.1) as shown in Fig. 2 was assessed in two different setups: The first setup consisted of a desktop connected to the EMTS and a webcam; while the second made use of an AR device (smartphone and HL2) containing a camera to locate the 3D skull and the



**Fig. 6** Data acquisition step

instrument wirelessly. The desktop configuration demonstrates the feasibility of the approach in general, whereas the smartphone/HL2 device implementation demonstrates that the approach is independent of the AR devices used (no device-based calibration). It also demonstrates the multi-device collaboration which can be relevant for surgical or training ends where multiple users can join the session. In both setups, the desktop and the AR devices used contained a tracking pattern module: Each device is responsible for detecting and tracking the 2D pattern in addition to rendering the model.

The following scenarios were considered: (1) moving instrument (i.e., pointer); (2) moving phantom skull; (3) moving marker; moving (1) and (2); moving (1), (2) and (3); in addition to the moving camera for the smartphone and HL2 setups.

### 3.7 User evaluation

A user evaluation was conducted to assess the final end-to-end output as experienced by different users using the OTS device (HL2).

#### 3.7.1 Task

To this end, 25 positions defined in the EMTS coordinate system were acquired and projected over the ground truth board. The acquisition included random slight rotations of the pointer. The calibration board was covered to hide any references of the divot points. Participants in this study had to align the pointer with the target augmentation (same rotation and position) using the HL2. The projection of position was represented by a cross section, whereas the orientation was represented by the yellow axis coming out toward the top of the board (Fig. 6).

#### 3.7.2 Volunteers

Ten volunteers were included in this user evaluation (6 males, 4 females) with different backgrounds (8 technical,

1 medical, 1 technical medicine). Their ages ranged between [20 – 29]( $n = 6$ ) and [30 – 39]( $n = 4$ ). Five participants reported that they were familiar with the use of the HL2.

### 3.7.3 Procedure

At the beginning of the experiment, volunteers were requested to fill a consent form which contains a description of the task. Standard eye calibration of the HL2 (Lewis et al. 2011) was performed by each participant, and a short explanation was provided. The goal of the eye calibration is to account for any display misalignment caused by the variance of inter-pupillary distance between the participants. Subsequently, the volunteers had to stand facing the multimodal marker in order to perform the alignment task as described in Sect. 3.7.1. Participants validated the correct alignment verbally by saying "Next" allowing the system to collect positional and orientation data. The system then showed the next adequate projection until the end of the experiment.

## 4 Results

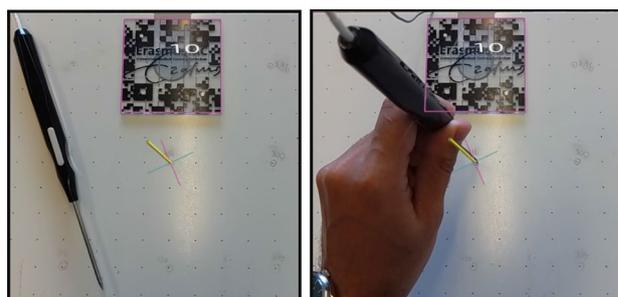
### 4.1 Pointer pose

The pointer's pose accuracy based on 18 distance measurements (100 mm ground truth intra-distances) had a mean average error of 0.28 mm 0.17mm. The minimum error is



**(a)** The projected registra- **(b)** Manual annotation of pin-  
tion centers using our solu- point spots  
tion

**Fig. 7** A pair of image before (left) and after (right) manual annotation



**(a)** The projected position **(b)** The alignment performed  
and orientation by a user

**Fig. 8** User evaluation: alignment task

equal to 0.06 mm, and the maximum one is equal to 0.60 mm.

### 4.2 Calibration assessment

The calibration error distribution for each scenario on the board, based on five calibrations for each one of them, is presented in Fig. 7. For each scenario, we notice a persistent error at the bottom right corner of the board, which we attribute to the tracking system's accuracy in that region. The images show a better accuracy around the multimodal marker, but the accuracy drops slightly on top of the marker (middle top of each heatmap for the middle and bottom row). Additionally, incrementing the number of calibration points and/or calibrating on a larger surface can improve the calibration accuracy. The latter observation can be confirmed by looking at each row of Fig. 7.

The average root mean square error (RMSE) over the five calibrations for each scenario is listed in 2). For the smallest surface area ( $60 \times 60$ ), the number of points used may have improved the calibration accuracy. For the larger surface areas ( $200 \times 200$  and  $400 \times 400$ ), this effect is hardly present. Thus, increasing calibration area as well as increasing

**Table 2** Calibration RMSE as a function of the calibration surface and the number of calibration points used

Surface (mm)	number of points	RMSE (mm)
60 × 60	4	2.03
	12	1.54
	24	1.41
200 × 200	4	1.16
	12	1.16
	24	1.28
400 × 400	4	1.21
	12	1.14
	24	1.10

the number of points may improve calibration accuracy. However, there is a limit to the improvement: For large areas, annotation (pointing out divots) errors have a smaller effect in estimating the orientation, while more points allow to average out the random errors in the annotations.

Figure 8 shows the mean RMSE for each of the individual calibrations. Each circle represents the error obtained from one of the five calibrations for a given surface, and their mean RMSEs are linked with a line. For the small surface area (60 × 60), we notice a larger variation between the five calibrations (std=0.45 mm) compared to calibration with large surface areas (200 × 200 and 400 × 400 mm with std=0.19 mm and 0.24 mm, respectively) when using four points for calibration. Figure 8 suggests that the use of 12 calibration points for the larger surface areas, or 24 points for all surface areas would guarantee a calibration error below 1.5 mm in similar conditions.

The calibration time, acquired while performing the calibrations, was 33 seconds for four points, 90 seconds for 12 points and finally, 195 seconds for 24 calibration point.

### 4.3 Marker tracking and instrument augmentation

The error distribution from the marker pose estimation and the multimodal-marker-calibration board assembly (no tracking system) had a mean error of 1.70 ± 0.50 [0.70-2.66] mm, whereas the pointer overlay error had a mean error of 2.50 ± 1.30 [0.80-6.26] mm.

Including the pointer in this assessment (thus including: calibration, tracking and pointer calibration) resulted in a mean error of 1.24 ± 1.05 [0.04-4.00] mm. The visualization of this contribution (Fig. 9 - difference) highlights the fact that compared to other regions of the calibration board, the central area yielded the smallest error contribution (< 2mm). A large error contribution is noticed on the borders of the board and more specifically the lower right corner.

### 4.4 Overlay accuracy assessment

Before overlaying the 3D model, the point-based image to patient registration (Horn and Berthold 1987) was performed five times on 22 registration spots to assess its variation. The image-to-patient registration had a mean fiducial registration error (FRE) of 2.16 ± 0.07 mm. The registration matrix used for the following experiments had a FRE of 1.60 mm.

The projection error estimated in mm based on the known diameter of the registration spots is shown in Table 3. For the monocular RGB camera, it is clear that the individual calibrations hardly affect the final accuracy, and thus that the calibration does not seem to be the major source of error. However, the projection angle is a quite relevant factor: The best results are obtained for 45° (2.37 mm for C3), and worse results were obtained for the 90° cases. (~ 10.5 mm).

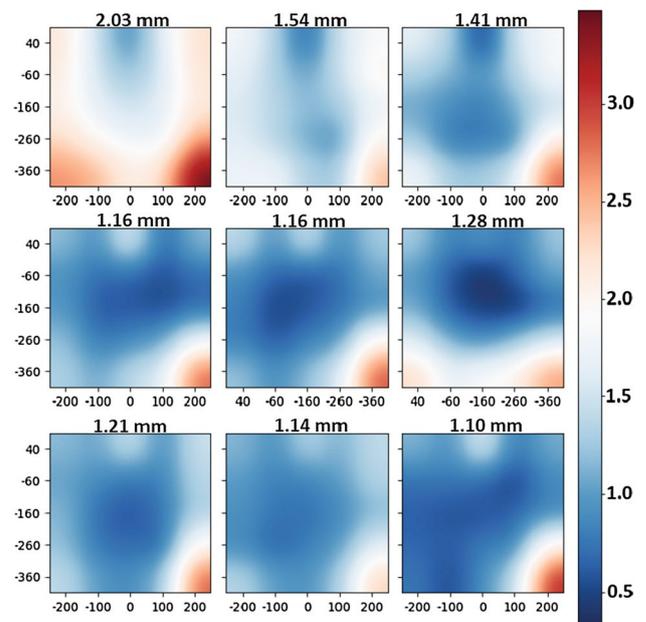


Fig. 9 The error distribution (Euclidean distance) for each calibration scenario on top of the assessment board. The multimodal marker is positioned as in Fig. 1). Each row represents the area covered by the calibration: top row: 60 × 60mm, middle row: 200 × 200mm and bottom row: 400 × 400mm. Each column represents the number of calibration points: 4, 12 and 24 from left to right

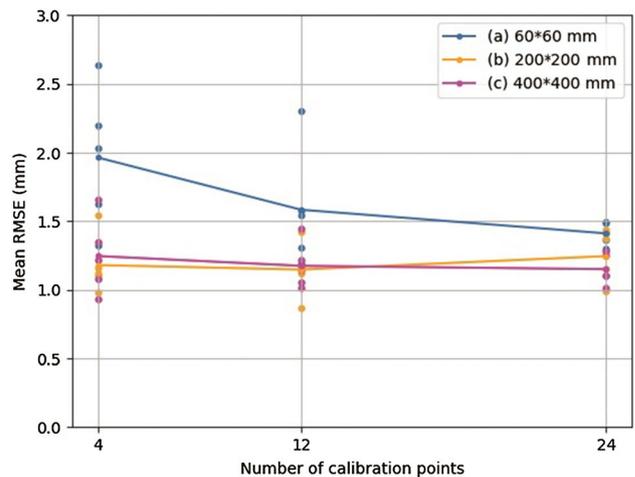
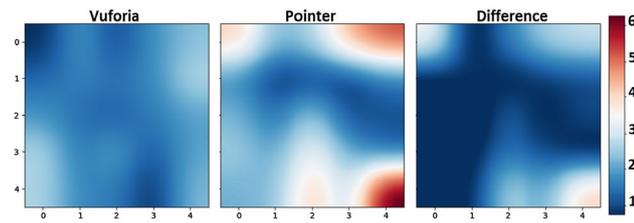


Fig. 10 The calibration error as a function of the number of calibration points

For the HL2, in contrast to the RGB camera, a difference between the calibrations can be observed:  $\epsilon$  values improved significantly (t-test p value < 0.05) from C1 (2.94 mm) to C2 (2.38 mm). However, no statistically significant improvement (p-value > 0.05) was obtained from C2 to C3 (2.16 mm). The standard deviations were less sensitive to the angle A of the device compared to the RGB results. In



**Fig. 11** Left: errors obtained from construction and marker pose estimation. Middle: errors obtained by overlaying a pointer tip. Right: the contribution of the tracking system, our calibration, and the provided pointer tip calibration

**Table 3** The overlay accuracy assessment results. The reported numbers are in function of the calibration chosen (C1, C2, C3) and the angle (A) of the camera compared to the position of the multimodal marker

Calibration	A	$\epsilon$ (mm)	RGB	$\epsilon$ (mm)	HL2
C1	20	2.77	0.36	2.91	0.62
	45	2.97	0.16	3.25	0.62
	90	10.06	1.33	2.67	0.13
C2	20	2.66	0.90	2.27	0.38
	45	2.69	1.21	2.55	0.43
	90	10.90	4.1	2.32	0.11
C3	20	2.54	0.80	2.20	0.50
	45	2.37	1.17	2.07	0.41
	90	10.40	3.94	2.21	0.17

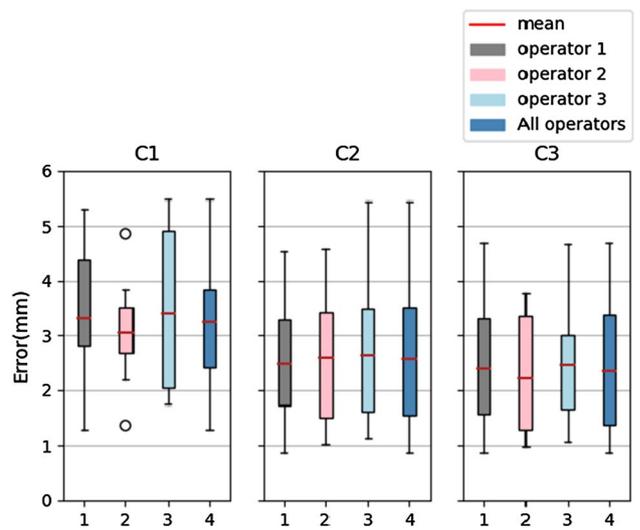
addition, the angle does not significantly affect the accuracy of the overlay for the HL2.

The results from the evaluation of the annotation-based overlay assessment are presented in Fig. 10.

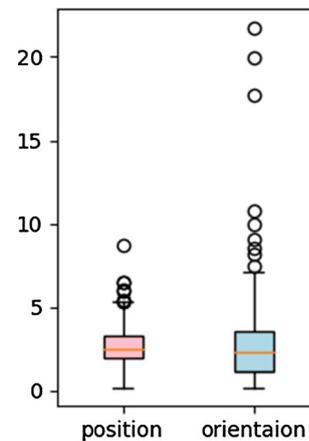
This figure shows that the annotators obtained similar trends (smaller error when using better calibrations). The mean shift in pixels obtained between the three annotators over a 120 annotated registration spot was 1.6 ± 0.8 pixel which in average equals to 0.5 ± 0.43 mm.

### 4.5 System output assessment

The configurations testing the visual output (Sect. 3.6) demonstrated the approach successfully. For all three configurations (desktop AR, smartphone AR, HL2), the phantom skull and the instrument could be augmented. The augmentations followed the objects in real-time for all the configurations. This demonstrates that with one calibration, devices that are capable of tracking the same multimodal marker, using the pattern tracking module (Fig. 2), can visualize the augmentations in multi-device AR collaboration. Figure 13 provides the visualizations obtained by both configurations simultaneously (smartphone and desktop), whereas Fig. 12 shows the projection of the skull, brain vessels, a tumour, and the



**Fig. 12** The results of the annotation performed by the three operators on 30 images for the three calibration scenarios C1, C2 and C3



**Fig. 13** Distribution of the volunteers’ performance in the alignment task

pointer which is placed behind the skull from the HL2’s point of view; One video demonstrating all the scenarios using the different devices (Sect. 3.6) in addition to some annotated pictures are enclosed in the supplementary material. In the HL2 part of the video, the recording mode and the automatic hand detection reduce the framerate. During regular use, without recording, the framerate is higher.

### 4.6 User evaluation

In total, 250 alignments (25 × 10) were performed. The results for the user evaluation are presented in Fig. 11. The volunteers achieved a mean positional error of 2.70 ± 1.28 mm, and a mean orientation error of 2.92 ± 2.90°.

## 5 Discussion

In this manuscript, an approach to align a device, such as a headset, smartphone, or tablet, containing an RGB camera with a surgical navigation system was proposed. This approach is generic and can be implemented both with electromagnetic and optical tracking systems. In our implementation, an EMTS was used. The approach was assessed in a crano-maxillofacial context. The described method could align the AR devices with the tracking system successfully, resulting in an augmented reality view where 3D models of the skull and the instrument (pointer) were projected on their physical counterparts.

Compared to previous methods (Zhu et al. 2017, 2018; Jiang et al. 2019), the approach presented in our study combines the technique which exclusively uses 2D patterns and the one where markers are attached to the device. The multimodal marker does not have to be fixated to an anatomical structure but can be moved around freely, as long as the 2D pattern is in the camera's view, and the attached sensor can be tracked by the tracking system. In contrast to earlier works (Kuzhagaliyev et al. 2018; Meulstee et al. 2019), the approach does not depend on a specific AR device or a navigation system's tracking technology. An example of a multimodal marker making use of an optical navigation system is shown in Fig. 14. To the best of our knowledge, no implementation aligning an electromagnetic tracking system with an AR device such as the HL has been explored before.

The calibration method proposed is to be performed only once: This means that our approach of integrating an AR device with a navigation system could facilitate collaborative augmented reality in the OR. Whether for learning purposes or (pre)intraoperative collaborations, any device that is equipped with a 2D pattern tracking module in addition to communication means such as Wi-Fi can receive the pose of the OR objects or the anatomical structures of interest.

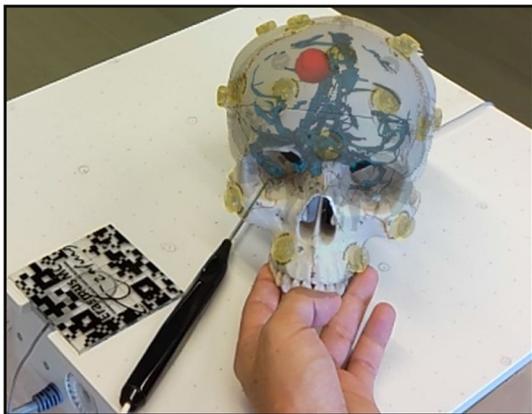


Fig. 14 Augmented view using the HoloLens 2

Moreover, only the headset and the multimodal marker, which can be prepared preoperatively, are added to the EM navigation setup; no changes were brought to the typical workflows when using electromagnetic navigation systems intraoperatively (Berger et al. 2015). This might help integrating AR into the OR smoothly (Figs. 15).

To summarize, the presented approach offers the following advantages over existing solutions: (1) AR-device independent; (2) one time calibration; (3) tracking system technology-independent (Fig. 16); (4) movable fiducial marker (detachable from patient anatomy); (5) capability of multi-device collaboration.

The calibration procedure is a prerequisite for the augmentation. We did not investigate the effect of marker size or shape on the final output in this study. The size of the 2D pattern was decided based on surgical space and a good



Fig. 15 Augmented images of the phantom skull and the pointer. The pictures show a collaborative augmented reality scenario where the blue skull is the desktop's view and the green skull is from the smartphone point of view. The left image showcases the instrument augmentation in both devices, while the right one represents the projections after moving the multimodal marker and the skull



Fig. 16 Multimodal marker using reflective spheres for an optical navigation system

initial pose estimation of the pattern. Assuming an  $80 \times 80$  mm pattern was too small for an accurate calibration, the calibration board was built as the pattern's extension. Its size was decided based on the necessary surgical volume and the tracking volume capabilities of the EMTS. The initial hypothesis assumed that a larger surface for calibration may lead to a smaller angular error. The results confirmed this hypothesis to some extent. Obviously, this claim is limited: For four points-based calibration, the alignment using the largest surfaces ( $200 \times 200$  and  $400 \times 400$  mm) performed best. However, when calibrating on a larger surface beyond  $200 \times 200$  mm, no substantial improvement is noticed. Additionally, adding more points to the largest surface area allowed a better representation of the calibration board when including inner divots; consequently, it led to a slightly better RMSE compared to the medium size calibration ( $200 \times 200$  mm). The calibration errors plots (Fig. 8) suggest that the accuracy cannot be improved beyond a certain point. It is limited by the pointer localization accuracy and the multimodal marker's fixation to the calibration board.

To assess the calibration accuracy's impact on the overlay, a point-based distance metric (in mm) was used for the 2D augmented images. There is a small difference in the overlay accuracy between the three calibrations. Specifically, for the camera-based solution, the accuracy changes are noticed on the  $20^\circ$  and  $45^\circ$  camera poses: The mean error for these poses slightly improved from C1 to C3. However, large errors were noticed for the  $90^\circ$  case. The reason behind this large shift is the missing RGB camera intrinsics from the Vuforia engine solution. This estimation is under-determined in case the marker is positioned at  $90^\circ$ , i.e., aligned with the camera viewing plane (as marker distance and camera viewing angle both determine marker size in that case). For the HL2-based solution, the standard deviations were similar for the different angles and much smaller compared to the results from the monocular camera case. The reason is that the camera intrinsics for HL2 are known to the Vuforia engine. It is important to mention that a better RMSE (TRE) on the calibration does not necessarily improve the overlay accuracy when the mean RMSE is less than a certain threshold. Even though C3 has a better RMSE compared to C2, the improvement in the overlay is not statistically significant.

The error heatmaps in Fig. 9 demonstrate that the contribution of the tracking error is generally small compared to Vuforia-only overlay except for the bottom right region of the calibration board, which consistently shows a larger error (even in the calibration heatmaps Fig. 7). This large error is attributed to the tracking system accuracy.

To measure the final output accuracy of the HL2, which includes the eye calibration error, a user

evaluation was conducted. The positional error (2.7 mm) was slightly larger than the annotation-based assessment of the skull projection (2.1 mm) (Table 3). However, orientation errors of the pointer were slightly higher and with more outliers. This may relate to how users perceive the alignment target when using OST devices to reach or match virtual objects (Singh et al. 2010; Swan et al. 2015). Many parameters such as color, opacity (Do et al. 2020; Ping et al. 2020), shape, size and pointer augmentation (Benmahdjoub et al. 2021) can impact the depth perception to improve or worsen the performance. To account for this, virtual mirrors have been suggested in the literature to view the target from different angles (Martin-Gomez et al. 2020, Alejandro et al. 2020). Traditionally, volunteers could also look at the target from various sides to perceive the spatial relationships better, and position the pointer at the right location. However, in our study, we did not enforce this on the participants.

The mean final output accuracy is 2.70 mm. It could be argued that such an error would not be tolerated for some surgical procedures which require a high accuracy. For instance, in mandibular-split osteotomy, a safe distance to the alveolar nerve could be as low as 2 mm. However, the system can still provide some benefits to surgeries like spring-assisted craniectomy, where the free-hand traditional techniques can reach an error up to 1cm of suture detection.

The current study presents a technical assessment including a user evaluation where an alignment task was performed. Clinical conditions need to be considered when integrating this system in a clinical scenario. Therefore, phantom studies will be conducted in the near future in order to determine strong and weak points of such an approach. Data on surgery outcome, the time taken during surgery and visualization efficiency for a specific application will be relevant for this goal.

## 6 Conclusion

In conclusion, we proposed and assessed an approach to align a surgical navigation system with an AR device which contains an RGB camera. The approach is demonstrated using an EMTS, a monocular RGB camera, a smartphone and Microsoft HL2. We successfully projected the 3D image of the skull on top of the physical one and evaluated the overall solution. The mean projection error on the image plane under three viewing angles using the best calibration was around 2.1 mm. Additionally, the user

evaluation performed by 10 users yielded a mean positional error of 2.7 mm using the HL2.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10055-022-00653-3>.

## Declarations

**Conflict of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article. The authors declare that no patient consent was required.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Azimi E, et al. (2017) Robust optical see-through head-mounted display calibration: taking anisotropic nature of user interaction errors into account. In: Proceedings—IEEE virtual reality, pp 219–220. <https://doi.org/10.1109/VR.2017.7892255>
- Azimi E, et al. (2020) An interactive mixed-reality platform for bedside surgical procedures. In: Medical image computing and computer assisted intervention - MICCAI 2020. Ed. by Anne L. Martel et al. Springer, Cham, pp 65–75. isbn: 978-3-030-59716-0
- Badiali G et al (2014) Augmented reality as an aid in maxillofacial surgery: validation of a wearable system allowing maxillary repositioning. *J Cranio-Maxillof Surg* 42(8):1970–1976. <https://doi.org/10.1016/j.jcms.2014.09.001>
- Banz VM et al (2016) Intraoperative image-guided navigation system: development and applicability in 65 patients undergoing liver surgery. *Langenbeck's Arch Surg* 401(4):495–502. <https://doi.org/10.1007/s00423-016-1417-0>
- Benmahdjoub M et al (2020) Augmented reality in craniomaxillofacial surgery: added value and proposed recommendations through a systematic review of the literature. *Int J Oral Maxillof Surg*. <https://doi.org/10.1016/j.ijom.2020.11.015>
- Benmahdjoub M et al (2021) Virtual extensions improve perception-based instrument alignment using optical see-through devices. *IEEE Trans Visual Comput Graph* 27(11):4332–4341. <https://doi.org/10.1109/TVCG.2021.3106506>
- Berger M et al (2015) Approach to intraoperative electromagnetic navigation in orthognathic surgery: a phantom skull based trial. *J Cranio-Maxillof Surg* 43(9):1731–1736. <https://doi.org/10.1016/J.JCMS.2015.08.022>
- Chen X et al (2015) Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display. *J Biomed Informatics* 55:124–131. <https://doi.org/10.1016/J.JBI.2015.04.003>
- Chotanaphuti T et al (2008) Comparative study between computer assisted-navigation and conventional technique in minimally invasive surgery total knee arthroplasty, prospective control study. *J Med Assoc Thailand Chotmaihet Thangphaet* 91:1382–8
- Cleary K, Terry MP (2010) Image-guided interventions: technology review and clinical applications. *Ann Rev Biomed Eng* 12(1):119–142. <https://doi.org/10.1146/annurev-bioeng-070909-105249>
- Do TD, Joseph JL, Ryan PM (2020) The effects of object shape, fidelity, color, and luminance on depth perception in handheld mobile augmented reality. In: Proceedings—2020 IEEE international symposium on mixed and augmented reality, ISMAR 2020, pp 64–72. [arXiv: 2008.05505](https://arxiv.org/abs/2008.05505)
- Eggers G, et al. (2009) Intraoperative computed tomography and automated registration for image-guided cranial surgery. In: Dentomaxillofacial Radiology 38.1. PMID: 19114421, pp. 28–33. <https://doi.org/10.1259/dmfr/26098099>
- Gavaghan K et al (2012) Evaluation of a portable image overlay projector for the visualisation of surgical navigation data: phantom studies. *Int J Comput Assist Radiol Surg* 7(4):547–556. <https://doi.org/10.1007/s11548-011-0660-7>
- Genc Y, Tuceryan M, Navab N (2002) Practical solutions for calibration of optical see-through devices. *Proc Int Symp Mixed Augment Real ISMAR 2002*:169–175. <https://doi.org/10.1109/ISMAR.2002.1115086>
- Gsaxner C, et al. (2019) Markerless image-to-face registration for untethered augmented reality in head and neck surgery. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Vol. 11768 LNCS. Springer, pp 236–244. ISBN: 9783030322533. [https://doi.org/10.1007/978-3-030-32254-0\\_27](https://doi.org/10.1007/978-3-030-32254-0_27)
- Hara T, et al. (2020) Efficacy of atlantoaxial transarticular screw fixation using navigation-guided drill: technical note. In: *World Neurosurgery* 134, pp 378–382. ISSN: 1878-8750. <https://doi.org/10.1016/j.wneu.2019.10.176>
- Horn BKP (1987) Closed-form solution of absolute orientation using unit quaternions. *J Opt Soc Am A* 4(4):629–642. <https://doi.org/10.1364/JOSAA.4.000629>
- Incekara F et al (2018) Clinical feasibility of a wearable mixed-reality device in neurosurgery. *World Neurosurg* 118:e422–e427. <https://doi.org/10.1016/j.wneu.2018.06.208>
- Itoh Y, Gudrun K (2014) Performance and sensitivity analysis of INDICA: INteraction-free display calibration for optical see-through head-mounted displays. In: ISMAR 2014—IEEE international symposium on mixed and augmented reality—science and technology 2014, proceedings, pp 171–176. <https://doi.org/10.1109/ISMAR.2014.6948424>
- Jiang T, et al. (2019) Precision of a novel craniofacial surgical navigation system based on augmented reality using an occlusal splint as a registration strategy. In: *Scientific reports* 9.1, p 501. ISSN: 20452322. <https://doi.org/10.1038/s41598-018-36457-2>
- Khenak N, Jeanne V, Patrick B (2020) Effectiveness of augmented reality guides for blind insertion tasks. In: *Frontiers in virtual reality* 2020.588217. ISSN: 2673-4192. <https://doi.org/10.3389/frvir.2020.588217>
- Kuzhagaliyev T, et al. (2018) Augmented reality needle ablation guidance tool for irreversible electroporation in the pancreas. In: *Medical imaging 2018: image-guided procedures, robotic interventions, and modeling*. Ed. by Robert J. Webster and Baowei Fei. Vol. 10576. SPIE, p. 30. isbn: 9781510616417. <https://doi.org/10.1117/12.2293671>
- Lewis JR, et al. (2011) Aligning interpupillary distance in a near-eye display system. 20130050642A1. url: <https://patents.google.com/patent/US20130050642A1/en>

- Li Y, et al. (2018) A wearable mixed-reality holographic computer for guiding external ventricular drain insertion at the bedside. *en. In: J Neurosurg*, p. 1. ISSN: 0022-3085. <https://doi.org/10.3171/2018.4.JNS18124>
- Makibuchi N, Haruhisa K, Akio Y (2013) Vision-based robust calibration for optical see-through head-mounted displays. In: 2013 IEEE international conference on image processing, ICIP 2013—proceedings, pp 2177–2181. <https://doi.org/10.1109/ICIP.2013.6738449>
- Martin-Gomez A, et al. (2020) Augmented mirrors. In: Proceedings—2020 IEEE international symposium on mixed and augmented reality, ISMAR 2020. Institute of Electrical and Electronics Engineers Inc., pp. 217–226. <https://doi.org/10.1109/ISMAR50242.2020.00045>
- Martin-Gomez A, et al. (2020) Gain a new perspective: towards exploring multi-view alignment in mixed reality. In: Proceedings —2020 IEEE international symposium on mixed and augmented reality, ISMAR 2020. Institute of Electrical 15 and Electronics Engineers Inc., pp. 207–216. <https://doi.org/10.1109/ISMAR50242.2020.00044>
- Meulstee JW et al (2019) Toward holographic-guided surgery. *Surg Innov* 26(1):86–94. <https://doi.org/10.1177/1553350618799552>
- MevisLab (2020). <https://www.mevislab.de/>. (Visited on 05/11/2020)
- Mezger U, Claudia J, Michael B (2013) Navigation in surgery. In: *Langen beck's archives of surgery* 398.4, pp 501–514. ISSN: 1435-2451. <https://doi.org/10.1007/s00423-013-1059-4>
- Ping J et al (2020) Effects of shading model and opacity on depth perception in optical see-through augmented reality. *J Soc Inf Disp* 28(11):892–904. <https://doi.org/10.1002/jsid.947>
- Singh G, et al. (2010) Depth judgment measures and occluding surfaces in near-field augmented reality. In: Proceedings—APGV 2010: symposium on applied perception in graphics and visualization, pp 149–156. isbn: 9781450302487. <https://doi.org/10.1145/1836248.1836277>
- Swan JE, Gurjot S, Stephen RE (2015) Matching and reaching depth judgments with real and augmented reality targets. *IEEE Trans Visual Comput Graph* 21(11):1289–1298. <https://doi.org/10.1109/TVCG.2015>
- Tuceryan M, Navab N(2000) Single point active alignment method (SPAAM) for optical see-through HMD calibration for AR. In: pp 149–158. <https://doi.org/10.1109/ISAR.2000.880938>
- Vuforia (2020). <https://developer.vuforia.com/downloads/sdk>. (Visited on 05/11/2020)
- Wang J et al (2014) Augmented reality navigation with automatic marker-free image registration using 3-d image overlay for dental surgery. *IEEE Trans Biomed Eng* 61(4):1295–1304. <https://doi.org/10.1109/TBME.2014.2301191>
- Westerfield G, Antonija M, Mark B (2015) Intelligent augmented reality training for motherboard assembly. *Int J Artif Intell Educ* 25(1):157–172. <https://doi.org/10.1007/s40593-014-0032-x>
- Zhu M et al (2017) A novel augmented reality system for displaying inferior alveolar nerve bundles in maxillofacial surgery. *Sci Rep* 7:42365. <https://doi.org/10.1038/srep42365>
- Zhu M et al (2018) Does intraoperative navigation improve the accuracy of mandibular angle osteotomy: comparison between augmented reality navigation, individualised templates and free-hand techniques. *J Plast Reconstruct Aesthet Surg* 71(8):1188–1195. <https://doi.org/10.1016/j.bjps.2018.03.018>
- Zinser MJ et al (2013) Computer-assisted orthognathic surgery: waferless maxillary positioning, versatility, and accuracy of an image-guided visualisation display. *Br J Oral Maxillof Surg* 51(8):827–833. <https://doi.org/10.1016/J.BJOMS.2013.06.014>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.