Computer Methods and Programs in Biomedicine 200 (2021) 105854

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Augmented Reality for Head and Neck Carcinoma Imaging: Description and Feasibility of an Instant Calibration, Markerless Approach

Christina Gsaxner[a,b,c,d,*], Antonio Pepe[a,c,d], Jianning Li[a,c,d], Una Ibrahimpasic[a,c], Jürgen Wallner[b,c,d,e,**], Dieter Schmalstieg[a,d], Jan Egger[a,b,c,d,**]

[a] Institute of Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16, 8010 Graz, Austria
[b] Department of Oral and Maxillofacial Surgery, Medical University of Graz, Auenbruggerplatz 5, 8036 Graz, Austria
[c] Computer Algorithms for Medicine Laboratory, Graz, Austria
[d] BioTechMed-Graz, Mozartgasse 12/II, 8010 Graz, Austria
[e] Department of Cranio-Maxillofacial Surgery, AZ Monica Hospital Antwerp and Antwerp University Hospital, Antwerp, Belgium

ARTICLE INFO

Article history:
Received 10 June 2020
Accepted 16 November 2020

Keywords:
Augmented Reality
Head and Neck Surgery
Computer-Assisted Surgery
Image-guided Surgery
Medical Visualization

ABSTRACT

Background and Objective: Augmented reality (AR) can help to overcome current limitations in computer assisted head and neck surgery by granting “X-ray vision” to physicians. Still, the acceptance of AR in clinical applications is limited by technical and clinical challenges. We aim to demonstrate the benefit of a marker-free, instant calibration AR system for head and neck cancer imaging, which we hypothesize to be acceptable and practical for clinical use.

Methods: We implemented a novel AR system for visualization of medical image data registered with the head or face of the patient prior to intervention. Our system allows the localization of head and neck carcinoma in relation to the outer anatomy. Our system does not require markers or stationary infrastructure, provides instant calibration and allows 2D and 3D multi-modal visualization for head and neck surgery planning via an AR head-mounted display. We evaluated our system in a pre-clinical user study with eleven medical experts.

Results: Medical experts rated our application with a system usability scale score of 74.8 ± 15.9, which signifies above average, good usability and clinical acceptance. An average of 12.7 ± 6.6 minutes of training time was needed by physicians, before they were able to navigate the application without assistance.

Conclusions: Our AR system is characterized by a slim and easy setup, short training time and high usability and acceptance. Therefore, it presents a promising, novel tool for visualizing head and neck cancer imaging and pre-surgical localization of target structures.

1. Introduction

In many clinical routines, medical imaging data, such as computed tomography (CT) or magnetic resonance imaging (MRI), is the major source of information for decision making. Computer-assisted intervention (CAI) aims to incorporate this information more naturally in the clinical workflow by using computerized tools and methods. CAI has greatly facilitated procedures in oral- and craniomaxillofacial surgery (OCMS), such as surgical planning, computer-assisted design, visualization, image-guided surgery (IGS) and others [1–3].

However, there are still shortcomings in current CAI procedures. Visualization and inspection of imaging data is usually still separated, both spatially and temporally, from the patient. Likewise, in IGS using conventional navigation systems, two-dimensional (2D) imaging is displayed on an external monitor. This results in a *switching focus* problem [4], since the physician is continuously forced to mentally map information between 2D imaging data and the 3D patient. Moreover, the usage of navigation systems in OCMS-IGS often relies on invasive, skull-fixed fiducials for image-to-patient registration [5], and usually involves lengthy set-up and calibration procedures which require time and human resources. One application of image-guided navigation is the pre-operative localization of the surgical target on the patient's outer surface [6]. The surgeon's interest in this examination lies on the topology of

* Corresponding author at: Institute of Computer Graphics and Vision, Graz University of Technology, Inffeldgasse 16, 8010 Graz, Austria.
** Corresponding authors at: Department of Oral and Maxillofacial Surgery, Medical University of Graz, Auenbruggerplatz 5, 8036 Graz, Austria.
E-mail addresses: christina.gsaxner@icg.tugraz.at (C. Gsaxner), j.wallner@medunigraz.at (J. Wallner), jan.egger@medunigraz.at (J. Egger).

https://doi.org/10.1016/j.cmpb.2020.105854
0169-2607/

the anatomical and pathological structures and size estimates in relation to the outer anatomy. For such a task, the overhead of setting up and calibrating a navigation system is often not justified.

Augmented reality (AR) has the potential to bridge these gaps by transferring radiological imaging data three dimensionally (3D) directly into the current clinical situation and onto the patient. CT, PET-CT or MRI scans can be visualized in 3D and in real time, effectively granting “X-ray vision” to the physician [7,8]. However, despite the obvious benefits, few AR systems have successfully been established in clinical use [9,10]. Aside from technical challenges, such as precision, latency and reliability of both registration and tracking, complicated hardware setups using bone-fixed markers and external tracking infrastructure, as well as time-intensive calibration procedures preclude adoption of AR within the clinic. Such systems are seen by medical staff as disruptive to their usual workflow, significantly impeding acceptance [11]. Nonetheless, for applications such as the pre-operative localization task, were submillimeter precision is not essential, simple AR systems can be a valuable visual aid, helping physicians to overcome the switching focus problem.

Current literature provides many experimental studies about the introduction of mobile AR into OCMS, as shown by a recent survey by Badiali et al. [5]. In example, Yamaguchi et al. [12] designed a wearable AR system for dental implant surgery. Even though they use a head-mounted display, their system requires stationary infrastructure for image-to-patient registration. Badiali et al. [13] use a video see-through head worn AR display for supporting maxillary re-positioning, relying on skull-fixed markers for registration and tracking. Zhu et al. [14,15], Lin et al. [16] and Zhou et al. [17] all developed AR systems for various OCMS applications with head-worn displays, focusing on designing a registration strategy based on a customarily designed, non-invasive visual markers, rigidly attached to an occlusial splint. Our group was the first to use the HoloLens (Microsoft Corporation, Redmond, WA, USA) for supporting OCMS for the immersive visualization of imaging data from head and neck cancer patients [18–21], with the main goal of enabling marker-free image-to-patient registration. Pietruski et al. [22] also used the HoloLens, in combination with an infrared-based tracking system, for AR-guided osteotomy in the cranio-maxillofacial area. It is evident that most of these studies focus on overcoming selected technical challenges, such as image-to-patient registration, while there are few reports about the integration and usability of such technologies into a clinical scenario. Therefore, our goal was to design an AR system which is easy to setup, learn and use within a clinical environment and furthermore assess it’s usability and accessibility within our target group, medical professionals working in the head and neck area.

Based on related work and our own experience, we identify three main aspects that determine the potential acceptance of AR in a clinical setting: (i) accuracy, reliability, (ii) usability, and (iii) perceptual quality of the visualization. With these considerations in mind, we designed a system which leverages AR for in situ visualization during inspection prior to head and neck cancer surgery. We use a commercial AR optical see-through head-mounted display (OST-HMD), the HoloLens, to deliver visualization of imaging data showing the location of the surgical targets inside the skull, registered with the patient. Our AR application performs automatic registration of medical data to the patient in the physical space without the need for manual calibration, invasive markers or stationary infrastructure. We achieve this by building the application on top of our previously introduced markerless, untethered registration scheme [21]. Our application runs solely on mobile hardware and can be set up in a matter of minutes, either bedside or in the operating room. Registration happens in seconds and continuously adjusts to the patient’s position, eliminating the need to sedate the patient. In this contribution, we describe our prototype implementation and evaluate our system in a pre-clinical study with 11 medical experts to gauge usability, benefit, and potential of AR in OCMS. Our hypothesis is that this form of markerless, instant calibration AR would be more acceptable and practical for simple tasks than high-end professional navigation systems, while, at the same time, granting novel, immersive insight in medical image data and the patient.

2. Materials and Methods

2.1. System overview and implementation

Our AR visualization system requires two components. For presenting virtual content in combination with the real environment, we use the HoloLens. It localizes itself within the environment and is equipped with a video camera and a time-of-flight depth sensor, which we utilize to capture the patient in both 2D and 3D. Since the computing power of the HoloLens is limited, computations requiring stronger hardware, such as face detection, surface reconstruction and point cloud registration, are outsourced to the second component, a PC or laptop.

Information from the HoloLens’ video camera, depth sensor and self-localization is streamed to the PC via WiFi and Transmission Control Protocol (TCP), were the accurate position of imaging data in relation to the physician is calculated. This pose information is transmitted back to the OST-HMD, where virtual content is placed accordingly. The visualization can be adjusted to the physician’s needs by using gaze and hand gestures as natively supported by the HoloLens, or, alternatively, a Bluetooth gamepad. All data is transmitted directly between HoloLens and device - an internet connection is not required.

This system architecture, were heavier computations are loaded off to a desktop PC and performed asynchronously, allows us to keep the HoloLens graphics processing unit unoccupied for rendering. Therefore, the application runs at near-to-full frame rate of approximately 50 - 60 frames per second (FPS) on the device. This is critical to ensure accurate, visually pleasing rendering and to minimize jitter and overall latency of the system. These factors are not only important for system accuracy, but also have a large impact on user comfort.

The HoloLens application was developed using Unity 3D, C# and the Microsoft Mixed Reality Toolkit, the client application in Python. An overview of the proposed system setup is shown in Fig. 1.

2.2. Medical data collection and model extraction

In an offline step, medical imaging data is collected from patients within the clinical routine. Our system works with CT, PET-CT and MRI scans routinely acquired in the treatment of oncology patients, without requiring a specialized scanning protocol. For 3D visualization, surface renderings of structures of interest, such as tumors and infiltrated bones, need to be extracted by segmenting anatomical structures and tissues of interest and converting them to polygonal meshes. We automatically segment bone and tumour using thresholding, and use the Marching Cubes algorithm for mesh decimation. The resulting meshes can be displayed on the HoloLens at real-time frame rates. Fig. 2 shows bone and tumor surfaces, registered with the patient, through the HMD. Furthermore, we extract a point cloud representation of the patient’s skin, which will be used by our system for automatic registration. Our application also supports 2D orthogonal multi-planar visualization for the exploration of volumetric medical data sets along all three standard anatomical orientations, as shown in Fig. 2 (c) and (d). 2D information can be loaded into our application in the form of individual images slices. These are interpolated and converted into a

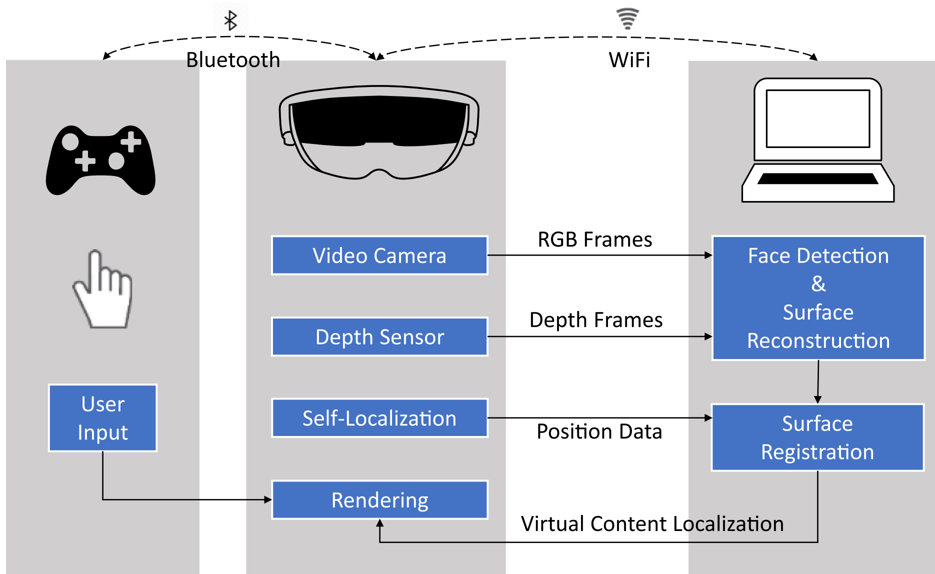

Fig. 1. System overview. We access the video camera and depth sensor of the HoloLens and annotate the frames with 3D positions from the built-in self-localization capability of the OST-HMD. RGB, depth and position are streamed over WiFi to the computer. Heavier computations, such as face detection and surface registration, which would not run fast enough on the HoloLens, are performed asynchronously on the PC. The HoloLens is only responsible for self-localization, tracking and rendering the graphics output.

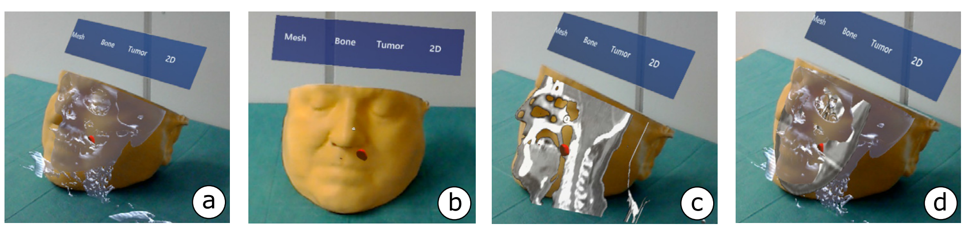

Fig. 2. Our application offers several modes of visualizing medical image data. Segmented surfaces can be rendered in 3D, as shown for the patient phantom's skull and a tumor (red) (a) and (b) for a tumor alone. Furthermore, the application can render orthogonal slices through the image volume in the three standard anatomical planes: (c) shows the sagittal orientation, in combination with the 3D rendered tumor, (d) the coronal orientation, combined with the tumor and skull in 3D.

3D texture, which is deployed to the HoloLens for seamless, multi-planar rendering.

2.3. Registration and calibration

Our AR application performs automatic registration of medical data to the patient in the physical space without the need for calibration, markers or external tracking systems. We base this pipeline on a previously introduced markerless, untethered registration scheme [21]. In summary, the registration principle exploits the fact that a 3D model of the patient anatomy can be created directly on the basis of pre-operative imaging. Furthermore, the information gathered by the sensors of the HoloLens is utilized to reconstruct the patient's face, as seen by the OST-HMD. In this way, anatomical imaging can automatically be matched to the patient by simply looking at them through the OST-HMD.

More specifically, we add frames arriving from the HoloLens into a buffer, and then draw the most recent matching RGB and depth image pairs based on their timestamps. We then use a convolutional neural network for real-time face detection of the patient on the RGB image. The single-shot multibox detector by Liu et al. [23] provides an adequate trade-off between accuracy, speed and ease of implementation for our use case. Once a region of interest around the patient's face is determined, we reconstruct a point cloud representation of it by applying an inverse perspective transformation on the depth pixels in the corresponding region. This point cloud is then registered with the point cloud obtained from pre-interventional imaging by adopting Fast Global Registration [17] for initialization and an Iterative Closest Point variant [24]. The registration algorithm can be tested with our online platform Studierfenster (www.studierfenster.at). After the two point clouds are aligned, their world-space localization can be determined by relating them to the corresponding position information from the HoloLens. The entire registration pipeline takes between 0.5 and 2 seconds to complete and dynamically adjust to the movement of the patient. Since it is performed asynchronously on a companion PC, this does not impede the rendering performance on the HMD itself, which continues to deliver near-to-full frame rates.

2.4. Clinical usage and workflow

To set up our system, physician only need to put on and power the HoloLens as well as the companion PC. After starting the application on both devices, navigation and control of our system is entirely performed within the AR environment. The physician is prompted to select the current patient from a drop-down list; see Fig. 3 (a). A box is displayed in the AR environment, guiding the user to look at the patient, as shown in Fig. 3 (b). When the patient's face is detected (Fig. 3 (c)), the patient-specific image data is registered to it. As soon as an adequate registration has been found, imaging data is displayed automatically, overlaid onto the patient.

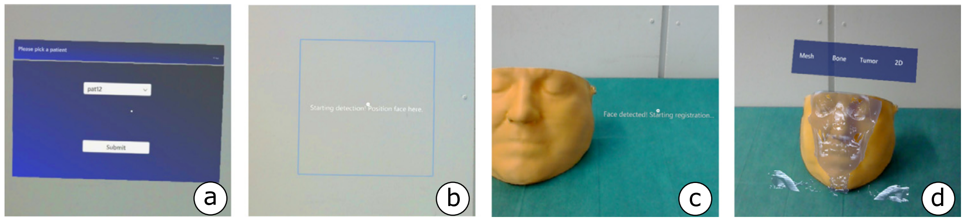

Fig. 3. The navigation through our system is entirely guided by augmented content. (a) The user selects the current patient and (b) is guided by a box to align the patient's face within the field of view of the HoloLens. (c) Once the patient's face is detected, the user is notified. (d) As soon as the registration process is finished, medical data is displayed overlaid with the patient. To interact with the scene, virtual buttons and sliders (top and right, respectively) can be used. Alternatively, the application may be controlled with a gamepad.

The physician can then toggle different visualization modes, including 3D surface structures and planar cuts. For multi-planar rendering, the user can seamlessly slice through the volume and change the anatomical orientation between axial, coronal and sagittal. In terms of input and navigation possibilities, the application can be controlled with virtual buttons or sliders (Fig. 3 (d)). These widgets can be operated using the HoloLens gaze and gesture input, or, alternatively, with a Bluetooth gamepad.

2.5. Pre-clinical study

The study was carried out in accordance with the local legal requirements and the Declaration of Helsinki (1975) at the Medical University of Graz, Austria and included the approval of the ethics committee of the university (30-340 ex 17/18, 31-416 ex 18/19). Informed consent was obtained from all subjects involved in this study.

Participants

We recruited eleven health care professionals with several years of clinical experience in practice and research in the head, neck and dental area at a university hospital. Six were specialised oral- and maxillofacial surgeons, two were neurosurgeons, two, dentists, and one was a senior experimental neuro-traumatologist. Four participants had previous experiences with OST-HMD devices; three also reported usage of AR-based systems in a medical context. The remaining seven participants did not report any exposure to OST-HMD devices or AR in the medical and non-medical domain.

Procedure

The study participants were first given an introduction to the HoloLens, including gaze and gesture interaction. The capabilities of the system and required workflow were demonstrated to them in a hands-on fashion using a patient phantom, which was 3D-printed from clinical PET-CT data. The data set and 3D model used for this phantom is publicly available [25]. This part of the study was labelled as training phase. Once the participants felt confident with using application, they were asked to test the system on a healthy human volunteer, who had an MRI scan done beforehand, in a testing phase. To simulate the clinical use case, a virtual tumor was placed within the head and neck area of this scan. After completing both phases of the study, the participants were asked to fill out a questionnaire collecting demographics and questionnaires about their impressions using the system in terms of usability, interface/interaction, visualization and overall impression. Finally, the participants were asked for verbal feedback in an informal interview. One study session, including training, testing, questionnaire and discussion, lasted approximately 35 minutes.

Measures

For each participant, we collected the times spent in training and testing phase. For measuring usability, we utilized the System Usability Scale (SUS) with ten items on a 5-point Likert scale [26]. Furthermore, we gave participants two questionnaires to rate the user interface/interaction and visualization on a 5-point Likert scale, representing a range of "strongly disagree" (1) to "strongly agree" (5). These survey statements, as well as a statistical analysis of the survey results, are given in Table 1.

Finally, users were asked to rate their overall impression from "very negative" (1) to "very positive" (5). Furthermore, most positive and most improvable aspects were determined in the form of essay questions.

Statistical analysis

The data gathered in this contribution is presented using descriptive (and analytical) statistical methods such as means ± standard deviations (SD). All statistical analyses were performed using Python with the open source package SciPy version 1.4.1 (https://www.scipy.org/).

3. Results

Training and testing times

In total, users spent an average of 12.7 ± 6.6 minutes in the training phase, and, 9.9 ± 4.9 minutes in the testing phase (see also Fig. 4). This results in a mean time of 22.6 ± 9.2 minutes of participants wearing the HMD. A two-sample t-test revealed that there was no significant difference between usage times of physicians experienced AR (averaging 25.3 ± 9.8 minutes) and users without prior experience (averaging 21.1 ± 9.2).

System usability

The perceived usability of the presented AR application was studied by a SUS questionnaire. On average, our participants rated the presented AR application with a SUS score of 74.8 ± 15.9, which, according to Brooke [26], means above average (> 68) usability of the system.

User interface, input and visualization

The results of the questionnaires are summarized as plots in Fig. 5; means and median values are shown in Table 1.

Overall impression and essay questions

Study participants had a very positive overall impression of the system, awarding an average of 4.5 ± 0.7 out of 5 points, with 1 representing "very negative" and 5 "very positive".

Concerning aspects of the system which stood out most positively, users mainly commented on the ease of use, intuitiveness and novel way of data representation. Among the aspects of the system which most urgently require improvement, most users mentioned the registration accuracy and time needed for auto-registration to converge.

4. Discussion

4.1. Accuracy and reliability

The implementation of medical AR systems faces major technical challenges, in particular, in registration and calibration [7]. Medical technology require the highest standards for reliability and are not tolerant of latency [27] or poor stability [28]. Occlusions

Table 1
Mean and standard deviation (SD), as well as median and inter-quartile range (IQR) of ratings given by users to the user interface, input and augmented reality visualization questionnaire. Statements were rated on a 5-point Likert scale representing a range of "Strongly disagree" (1) to "Strongly agree" (5).

Survey statement	Mean	SD	Median	IQR
Q1: Controlling the system was easy and intuitive.	4.09	0.83	4	1.75
Q2: The system provides enough guidance and feedback.	3.81	1.17	4	2
Q3: The user interface is well arranged and has an appealing style.	4.18	0.75	4	1
Q4: User input using gaze and air tap gestures was easy and intuitive.	3.54	1.21	3	2
Q5: User input using the gamepad was easy and intuitive.	4.36	0.81	5	1
Q6: The response time of the system to user input was satisfying.	4.00	1.34	4	1
Q7: I am satisfied with the registration time of the system.	3.54	1.36	3	2
Q8: The virtual content was accurately overlaid with the real world.	3.18	0.98	3	1
Q9: I am satisfied with the medical content presented in the AR environment.	4.27	1.01	5	1
Q10: I am satisfied with *how* the medical content is presented in the AR environment.	4.09	0.70	4	1
Q11: I believe the visualization of medical data in AR has benefits over the traditional method.	4.63	0.50	5	1

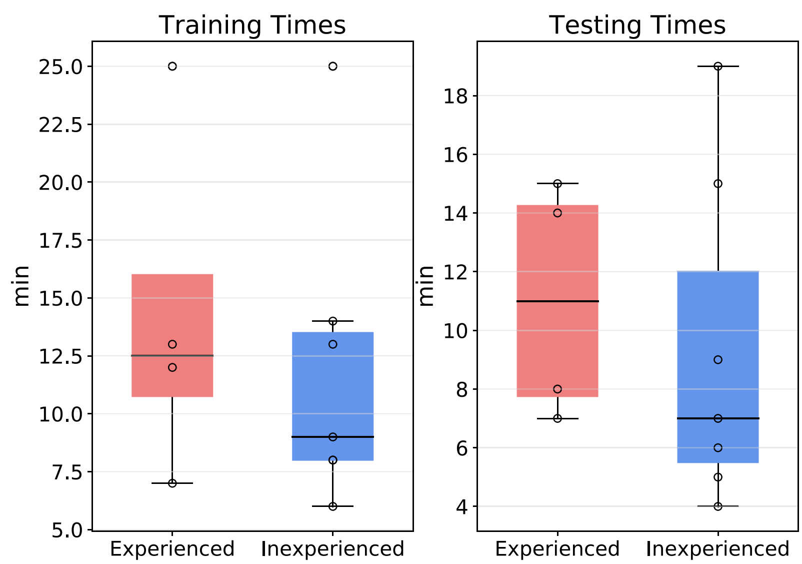

Fig. 4. Distribution of training and testing times (in minutes) for users with and without prior experience with augmented reality. The black bold line indicates the median, while the boxes show the interquartile range (IQR) between 25th and 75th percentile and whiskers indicate minimum and maximum values within 1.5 × IQR.

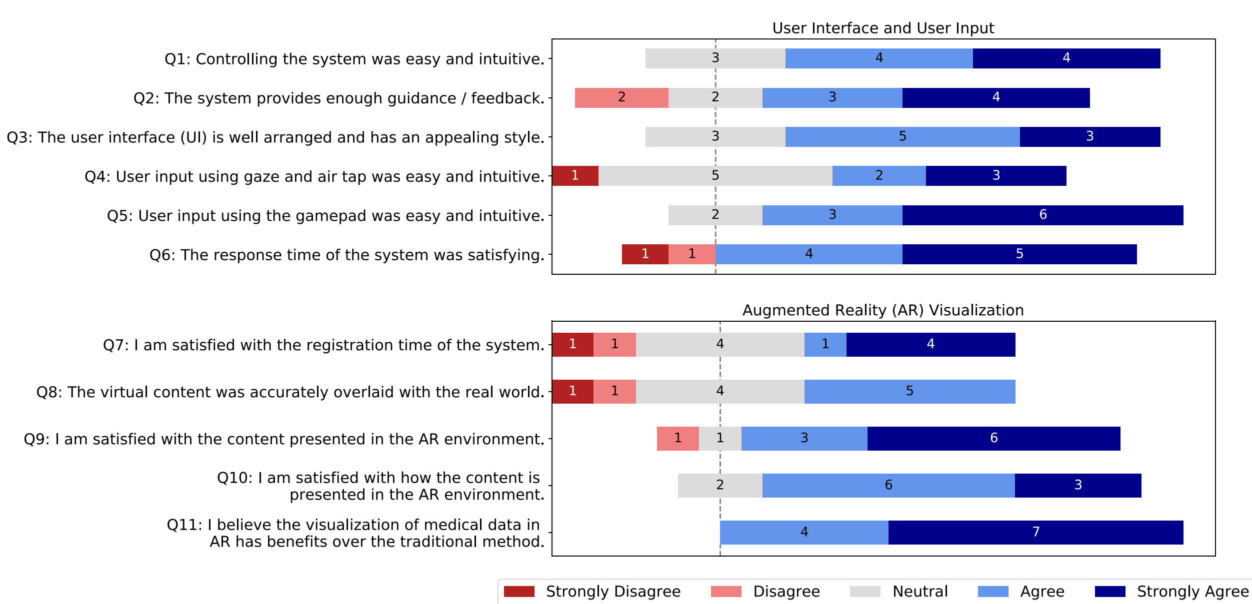

Fig. 5. Summary of questionnaire responses concerning user interface, user input and augmented reality (AR) visualization. Statements were rated on a 5-point Likert scale, ranging from "Strongly disagree" to "Strongly agree".

from instruments or blood, reflection on tissues or the unintentional misplacement of markers during intervention may interfere with registration. Unnoticed failures may misguide the physician, leading to a possibly serious medical accident [10]. We address this challenge by combining state-of-the-art technology with a field of application that is relatively tolerant of small amounts of error, namely pre-surgical inspection, visualization and localization of target structures.

Our study revealed deficiencies in terms of registration time and accuracy, with most participants commenting on a noticeable misalignment between virtual objects and the real patient, quantifying the error with estimates between a few millimeters of up to two centimeters. Those shortcomings can, at least partly, be attributed to the limitations of the chosen hardware. In a previous study, the accuracy of the employed image-to-patient registration was quantified with an average target registration error between patient and virtual content of 9.2 mm; the average error in translation was 3.9 mm and the average error in rotation of 4.9° [21]. This overall system error is caused by several sources. Our registration approach is dependent on the HoloLens' built in SLAM system, which has been shown to be subject to many environmental influences in a recent paper [29]. Furthermore, depth data acquired from the headset is of rather low quality and resolution, resulting in a poor representation of the patient's face after point cloud reconstruction. Also, surface reconstruction from CT scans involves some approximations, contributing to the overall error. Another source of inaccuracy comes from the point cloud registration algorithm itself, which is working stably and robustly, but of course not without a residual error. Aside from that, limited computing power of the headset prevents costly computations on the HoloLens itself, and therefore accounts for the need of a streaming-based approach, which obviously introduces additional latency. On top of that, the quantitative registration error might not be equivalent to the perceived misalignment by the user, but instead present a lower bound. Other factors, in particular drift or flickering of augmented content in the operator's view and incorrect display calibration deteriorate the end-to-end alignment accuracy.

Concerning registration time, our registration pipeline takes between 0.5 and 2 seconds to converge. Even though the visualisation dynamically adjusts if the patient moves, movement is, consequently, not detected in real-time frame rates. This causes some latency until the virtual content re-adjusts to its new position.

Our approach is, different from commercial navigation systems or related prototypes, markerless and calibration-free, but performs below the accuracy of marker-based, high-precision infrared tracking, as used by many of those solutions. We expect that, with the availability of more capable mobile hardware, these issues can be overcome without too much difficulty.

4.2. Usability

Introducing an AR system into a clinical workflow requires the addition of several components into an already complex environment. This includes hardware components as well as new or adapted work steps. To allow a smooth integration into clinical practise, none of these components should be seen by physicians as excessively time consuming or disruptive [10]. Thus, our AR system only requires a slim setup and a minimal amount of technical knowledge, to enable accessibility for a wide variety of medical professionals. It is easy to navigate and the display is not overloaded with information, such that the physician is not distracted.

In our user study, we found no significant difference in training, testing and total usage times of the presented system between users with and without experience in AR and HMDs. Therefore, we conclude that our system is easy to learn and to apply, even for people without prior experience.

With a SUS score of 74.8, our user study revealed above average usability of the presented visualization application. According to the adjective SUS rating system introduced by Bangor, Kortum and Miller [30], this constitutes a “Good” (5) system, on a 7-point scale ranging from “Worst imaginable” (1) to “Best imaginable” (7). A score above 70 also implies that the system would generally be accepted by the target audience. Therefore, we conclude that it is accessible to medical professionals from different backgrounds.

The questionnaire regarding user interface and input revealed that users found our system intuitive and easy to control, especially with the gamepad as interaction interface. Per contra, it has to be pointed out that this interaction paradigm has the drawback of keeping the hands of physicians occupied, and means that an additional component has to be introduced, which, in a clinical environment, always raises concerns with respect to hygiene and sterility. The alternative gaze and gesture-based interaction, however, was a source of frustration for participants. This interaction paradigm, which is native to the HoloLens, turned out to be difficult for users to perform, especially when interacting with smaller virtual objects. However, most users grew accustomed to the gestural input after an initial adjustment period, and rated it as acceptable. We anticipate that future hardware will make accessory-free input more user-friendly.

Users were furthermore impressed with the comfort of wearing the HMD. Despite the quite bulky design, they felt no substantial discomfort and were not bothered by the OST-HMD's weight. It has to be mentioned, however, that none of the participants wore the OST-HMD for more than 40 minutes.

4.3. AR visualization

The overarching goal of medical AR is to improve perception of the imaging data in support of the physician's decision-making process [9]. Oral- and craniomaxillofacial surgeons rely on medical imaging, such as CT, PET-CT or MRI, presented on 2D monitors, for pre-operative oro-pharyngeal tumor staging, tumor visualization and surgical treatment planning [31]. In our system, we preserve the orthogonal slice view physicians are accustomed to, and enhance it by adding structures of interest displayed in real 3D through the OST-HMD. Furthermore, imaging data and the patient can be examined at the same time, in the same physical reference frame.

All participants agreed that AR visualization has benefits over the traditional visualization of medical data on a 2D monitor. This can be mainly related to the fact that the investigated system provides 3D clinical imaging directly on the patient without using any external displays. Our study findings highlight the need for visualizing medical content appropriately for the current use case, especially in AR environments. For our clinical scenario, the shape and extent of the tumor, as well as its spatial localization with respect to bone, were of main interest to the surgeons. Consequently, users pointed out that the visualization of the tumor as 3D surface model is very informative, since the shape and extent is immediately available and need not to be mentally inferred by looking at planar slices from multiple orientations. However, for assessing the spatial relationship of a tumoral mass to the surrounding tissue, participants appreciated a 2D planar view of CT data. This can be explained by the fact that, by visualizing content on a plane, one degree of freedom is fixed, and distances in this plane are easier to estimate [7,32].

However, participants were not fully satisfied with the AR visualization, with incorrect depth perception being one the main issues, which is a known problem when using AR technology in a clinical scenario [33]. Depth perception in AR suffers from the fact that a naive overlay of virtual content over the real scene does not result in any depth cues, based on which the human perceptual system estimates ordering of objects and object distances. This is especially problematic for X-ray visualization, were virtual content should appear to lie within real objects. Users said that, even though the registration of virtual content to the patient was correct, content was appearing "*on top of, not within the patient*". More sophisticated visualization techniques, such as focus and context visualization (F+C) or ghosting exist to address this problem [34,35]. For example, Bichlmeier et al. [36] and Lerotic et al. [37] used F+C visualization for medical AR, which allows a non-restricted view of the target anatomy by concealing occluding anatomy, but integrating prominent features of the occluder with the virtual content for more intuitive perception. Hansen et al. [38] utilized illustrative rendering to create a sense of perspective in AR However, all these strategies were designed for video see-through displays, were real and virtual pixels can be manipulated. An implementation of such methods for an OST-HMD, were only the virtual content can be adjusted, is not straight forward and, unfortunately, research in this direction has stalled in the last years.

Another perceptual problem arises from the fact that current AR devices, such as the HoloLens, are not designed for supporting medical procedures. The HoloLens renders content on a fixed virtual display at an optical distance of around 2 metres away from the user, to which the user's eyes adjust their focus (or, accommodate). It is therefore recommendable to place virtual objects at around this distance, e.g. Microsoft recommends an ideal zone of 1.25 to 5 metres [39]. However, this distance is not feasible for medical or surgical applications, where the patient is typically in arm's reach of the patient, at distances between approximately 0.4 and 1.5 metres. Placing objects this near to the user leads to a convergence of the eyes to a distance different from the accommodation distance, resulting in a so-called vergence-accommodation conflict. This conflict can cause visual discomfort, fatigue or even decreased depth perception for the user [40]. In our study, this issue might not have been of great significance, since users only wore the device for a relatively short amount of time. However, for an application during an actual surgery, which might require the surgeon to wear the device for several hours, this will become a serious issue. Again, this issue cannot be resolved through software design alone, but requires specialized hardware solutions.

4.4. Limitations of the study

The pre-clinical nature of the presented study is one of it's main limitations. All results were obtained in a laboratory setting, in ideal conditions and without interference which might occur in a clinical environment. However, preliminary clinical results of similar technology published in current literature demonstrate the clinical applicability of AR and the possibility of translation of this technology into clinical scenarios [41–44]. We focus our study on acceptance and usability by medical experts, a factor which is usually neglected by related investigations. Nonetheless, a future clinical study is needed to ensure a measurable clinical benefit of the proposed system to patients.

4.5. Future outlook

Currently, we focus our application on pre-surgical scenarios, such as pre-operative localization of target structures and immersive visualization of imaging data, which may aid physicians during planning a surgery. For usage during actual surgery, several issues have to be overcome.

Some of these issues can only be partly addressed by software design, instead, OST-HMD technology itself has to be further developed on a technical and hardware basis to overcome them. In example, the overall registration error of our pipeline does not satisfy the millimetre precision requirement of IGS procedures. Limitations inherent to the HoloLens headset, such as instabilities in the SLAM system and low quality depth data, have a large impact on this error. We expect that, with the availability of more capable mobile hardware, these issues can be overcome without too much difficulty. Concerning the visualization of virtual content, current OST AR displays do not have the capabilities to produce realistic depth cues for near-interaction medical applications or true X-ray visualization, which causes both incorrect depth perception and visual discomfort. Current OST-HMDs, such as the HoloLens, are unfortunately not designed for medical purposes and, therefore, many interesting clinical applications, such as IGS, cannot satisfyingly be addressed with them. AR devices specifically designed for the medical sector are, therefore, highly desirable.

Moreover, during a maxillofacial surgery, the head and parts of the face may be moved, the skin might be opened and other parts of the face covered with surgical drapes. For IGS, our surface-based registration could only be used for an initial image-to-patient registration. Either the patient then remains stationary, e.g. by fixing his head with a frame, or an additional tracking strategy, which is more robust to such perturbations, must be employed.

5. Conclusion

In this study, we investigated the utility and usability of an AR visualization system for pre-operative localization and visualization in OCMS. We described a more simple surgical navigation approach using instant calibration, without the requirement of time-intensive, complex setups or markers. The AR application provides automatic registration of diagnostic imaging data to the patient's face and several modes of clinical data visualization. The pre-clinical user study found that, overall, physicians attributed high usability and acceptance to our system, rating it as easy to set up, learn and work with. The results indicate that AR improves the perception of medical data, and appropriate visualization aids the physicians in their decision-making process. As an upcoming project, we plan a series of clinical tests of our system with head and neck oncology patients, to further advance the translation of AR into clinical practice.

Role of the Funding Source

The funding sources had no involvement in the study design, collection, analysis and interpretation of data, writing of the report or decision to submit the article for publication.

Disclosure of conflicts of interest

Authors declare no conflict of interest.

Acknowledgment

Funding: This work was supported by the Austrian Science Fund (FWF) KLI 678-B31 "enFaced", the Austrian Research Promotion Agency (FFG) COMET K-Project 871132 "CAMed", and the TU Graz LEAD Project "Mechanics, Modeling and Simulation of Aortic Dissection". We also acknowledge BioTechMed-Graz.

References

[1] A.D. Nijmeh, N.M. Goodger, D. Hawkes, P.J. Edwards, M. McGurk, Image-guided navigation in oral and maxillofacial surgery, British Journal of Oral and Maxillofacial Surgery 43 (4) (2005) 294–302.
[2] M. Rana, H. Essig, A.M. Eckardt, F. Tavassol, M. Ruecker, A. Schramm, N.-C. Gellrich, Advances and innovations in computer-assisted head and neck oncologic surgery, Journal of Craniofacial Surgery 23 (1) (2012) 272–278.

[3] K.A. Rodby, S. Turin, R.J. Jacobs, J.F. Cruz, V.J. Hassid, A. Kolokythas, A.K. Antony, Advances in oncologic head and neck reconstruction: systematic review and future considerations of virtual surgical planning and computer aided design/computer aided modeling, Journal of Plastic, Reconstructive & Aesthetic Surgery 67 (9) (2014) 1171–1185.
[4] J.W. Meulstee, J. Nijsink, R. Schreurs, L.M. Verhamme, T. Xi, H.H.K. Delye, W.A. Borstlap, T.J.J. Maal, Toward holographic-guided surgery, Surgical innovation 26 (1) (2019) 86–94.
[5] G. Badiali, L. Cercenelli, S. Battaglia, E. Marcelli, C. Marchetti, V. Ferrari, F. Cutolo, Review on Augmented Reality in Oral and Cranio-Maxillofacial Surgery: towards “surgery-specific” head-up displays, IEEE Access 8 (2020) 59015–59028.
[6] F. Incekara, M. Smits, C. Dirven, A. Vincent, Clinical Feasibility of a Wearable Mixed-Reality Device in Neurosurgery, World Neurosurgery 118 (2018) e422–e427.
[7] T. Sielhorst, M. Feuerstein, N. Navab, Advanced medical displays: A literature review of augmented reality, Journal of Display Technology 4 (4) (2008) 451–467.
[8] X. Chen, L. Xu, Y. Wang, H. Wang, F. Wang, X. Zeng, Q. Wang, J. Egger, et al., Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display, Journal of Biomedical Informatics 55 (2015) 124–131.
[9] C.A. Linte, K.P. Davenport, K. Cleary, C. Peters, K.G. Vosburgh, N. Navab, P. Jannin, T.M. Peters, D.R. Holmes III, R.A. Robb, et al., On mixed reality environments for minimally invasive therapy guidance: systems architecture, successes and challenges in their implementation from laboratory to clinic, Computerized Medical Imaging and Graphics 37 (2) (2013) 83–97.
[10] L. Chen, T.W. Day, W. Tang, N.W. John, Recent developments and future challenges in medical mixed reality, in: Proc. International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, 2017, pp. 123–135.
[11] P. Markelj, D. Tomaževič, B. Likar, F. Pernuš, A review of 3D/2D registration methods for image-guided interventions, Medical image analysis 16 (3) (2012) 642–661.
[12] S. Yamaguchi, T. Ohtani, H. Yatani, T. Sohmura, et al., Augmented reality system for dental implant surgery, in: Proc. International Conference on Virtual and Mixed Reality (VMR), Springer, 2009, pp. 633–638.
[13] G. Badiali, V. Ferrari, F. Cutolo, C. Freschi, D. Caramella, A. Bianchi, C. Marchetti, et al., Augmented reality as an aid in maxillofacial surgery: Validation of a wearable system allowing maxillary repositioning, Journal of Cranio-Maxillofacial Surgery 42 (8) (2014) 1970–1976.
[14] M. Zhu, G. Chai, L. Lin, Y. Xin, A. Tan, M. Bogari, Y. Zhang, Q. Li, Effectiveness of a novel augmented reality-based navigation system in treatment of orbital hypertelorism, Annals of Plastic Surgery 77 (6) (2016) 662–668.
[15] M. Zhu, F. Liu, G. Chai, J.J. Pan, T. Jiang, L. Lin, Y. Xin, Y. Zhang, Q. Li, et al., A novel augmented reality system for displaying inferior alveolar nerve bundles in maxillofacial surgery, Scientific Reports 7 (2017) 42365.
[16] L. Lin, Y. Shi, A. Tan, M. Bogari, M. Zhu, Y. Xin, H. Xu, Y. Zhang, L. Xie, G. Chai, et al., Mandibular angle split osteotomy based on a novel augmented reality navigation using specialized robot-assisted arms - A feasibility study, Journal of Cranio-Maxillofacial Surgery 44 (2) (2016) 215–223.
[17] Q.Y. Zhou, J. Park, V. Koltun, Fast global registration, in: Proc. European Conference on Computer Vision (ECCV), Springer, 2016, pp. 766–782.
[18] A. Pepe, G.F. Trotta, C. Gsaxner, J. Wallner, J. Egger, D. Schmalstieg, V. Bevilacqua, et al., Pattern Recognition and Mixed Reality for Computer-Aided Maxillofacial Surgery and Oncological Assessment, in: Proc. Biomedical Engineering International Conference (BMEiCON), IEEE, 2019, pp. 1–5.
[19] A. Pepe, G.F. Trotta, P. Mohr-Ziak, C. Gsaxner, J. Wallner, V. Bevilacqua, J. Egger, et al., A Marker-Less Registration Approach for Mixed Reality-Aided Maxillofacial Surgery: a Pilot Evaluation, Journal of Digital Imaging 32 (6) (2019) 1008–1018.
[20] M. Sylos Labini, C. Gsaxner, A. Pepe, J. Wallner, J. Egger, V. Bevilacqua, et al., Depth-Awareness in a System for Mixed-Reality Aided Surgical Procedures, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 11645 LNAI, Springer Verlag, 2019, pp. 716–726.
[21] C. Gsaxner, A. Pepe, J. Wallner, D. Schmalstieg, J. Egger, Markerless image-to–face registration for untethered augmented reality in head and neck surgery, in: Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2019, pp. 236–244.
[22] P. Pietruski, M. Majak, E. Światek-Najwer, Z. Magdalena, M. Popek, M. Mazurek, M. Świecka, J. Jaworowski, Supporting mandibular resection with intraoperative navigation utilizing augmented reality technology - A proof of concept study, Journal of Cranio-Maxillofacial Surgery 47 (6) (2019) 854–859.
[23] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A.C. Berg, Ssd: Single shot multibox detector, in: Proc. European Conference on Computer Vision (ECCV), Springer, 2016, pp. 21–37.
[24] S. Rusinkiewicz, M. Levoy, Efficient Variants of the ICP Algorithm, in: Proc. International Conference on 3-D Digital Imaging and Modeling, IEEE, 2001, pp. 145–153.
[25] C. Gsaxner, J. Wallner, X. Chen, W. Zemann, J. Egger, Facial model collection for medical augmented reality in oncologic cranio-maxillofacial surgery, Scientific data 6 (1) (2019) 310.
[26] J. Brooke, Sus-a quick and dirty usability scale, Usability evaluation in industry 189 (194) (1996) 4–7.
[27] M.E. de Oliveira, H.G. Debarba, A. Lädermann, S. Chagué, C. Charbonnier, A hand-eye calibration method for augmented reality applied to computer-assisted orthopedic surgery, The International Journal of Medical Robotics and Computer Assisted Surgery 15 (2) (2018) e1969.
[28] R. Vassallo, A. Rankin, E.C.S. Chen, T.M. Peters, Hologram stability evaluation for Microsoft HoloLens, in: Proc. SPIE Medical Imaging, volume 10136, International Society for Optics and Photonics, 2017, p. 1013614.
[29] M.E. de Oliveira, H.G. Debarba, A. Lädermann, S. Chagué, C. Charbonnier, et al., A hand-eye calibration method for augmented reality applied to computer-assisted orthopedic surgery, The International Journal of Medical Robotics and Computer Assisted Surgery 15 (2) (2019) e1969.
[30] A. Bangor, P. Kortum, J. Miller, Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale, Journal of Usability Studies 4 (3) (2009) 114–123.
[31] A. Mahajan, A. Ahuja, N. Sable, H.E. Stambuk, et al., Imaging in oral cancers: A comprehensive overview of imaging findings for staging and treatment planning, Oral Oncology 104 (2020) 104658.
[32] J. Traub, P. Stefan, S.M. Heining, T. Sielhorst, C. Riquarts, E. Euler, N. Navab, Hybrid navigation interface for orthopedic and trauma surgery, in: Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2006, pp. 373–380.
[33] M.A. Livingston, A. Dey, C. Sandor, B.H. Thomas, Pursuit of 'X-ray vision' for augmented reality, in: Human Factors in Augmented Reality Environments, Springer, 2013, pp. 67–107.
[34] J. Kruger, J. Schneider, R. Westermann, Clearview: An interactive context preserving hotspot visualization technique, Transactions on Visualization and Computer Graphics 12 (5) (2006) 941–948.
[35] D. Kalkofen, E. Mendez, D. Schmalstieg, Interactive focus and context visualization for augmented reality, in: Proc. International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, 2007, pp. 191–201.
[36] C. Bichlmeier, F. Wimmer, S.M. Heining, N. Navab, et al., Contextual anatomic mimesis: Hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality, in: Proc. International Symposium on Mixed and Augmented Reality (ISMAR), 2007, pp. 129–138.
[37] M. Lerotic, A.J. Chung, G. Mylonas, G.Z. Yang, et al., Pq-space based non-photorealistic rendering for augmented reality, in: Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, 2007, pp. 102–109.
[38] C. Hansen, J. Wieferich, F. Ritter, C. Rieder, H.O. Peitgen, et al., Illustrative visualization of 3D planning models for augmented reality in liver surgery, International Journal of Computer Assisted Radiology and Surgery 5 (2) (2010) 133–141.
[39] Microsoft, Mixed Reality - Comfort, 2020. https://docs.microsoft.com/en-us/windows/mixed-reality/comfort.
[40] D.M. Hoffman, A.R. Girshick, K. Akeley, M.S. Banks, et al., Vergence-accommodation conflicts hinder visual performance and cause visual fatigue, Journal of Vision 8 (3) (2008). 33–33
[41] Y. Li, X. Chen, N. Wang, W. Zhang, D. Li, L. Zhang, X. Qu, W. Cheng, Y. Xu, W. Chen, Q. Yang, A wearable mixed-reality holographic computer for guiding external ventricular drain insertion at the bedside, Journal of Neurosurgery 131 (5) (2019) 1599–1606.
[42] T.P.C. Van Doormaal, J.A.M. Van Doormaal, T. Mensink, Clinical accuracy of holographic navigation using point-based registration on augmented-reality glasses, Operative Neurosurgery 17 (6) (2019) 588–593.
[43] W. Han, X. Yang, S. Wu, S. Fan, X. Chen, Z.M. Aung, T. Liu, Y. Zhang, S. Gu, G. Chai, et al., A new method for cranial vault reconstruction: Augmented reality in synostotic plagiocephaly surgery, Journal of Cranio-Maxillofacial Surgery 47 (8) (2019) 1280–1284.
[44] B. Carl, M. Bopp, B. Voellger, B. Saß, C. Nimsky, et al., Augmented Reality in Transsphenoidal Surgery, World Neurosurgery 125 (2019). e873–e883