Otology & Neurotology
39:e1137–e1142 © 2018, Otology & Neurotology, Inc.

Development of a Mixed Reality Platform for Lateral Skull Base Anatomy

*Jonathan L. McJunkin, *Pawina Jiramongkolchai, †Woenho Chung, †Michael Southworth, *Nedim Durakovic, *Craig A. Buchman, and †Jonathan R. Silva

**Department of Otolaryngology—Head and Neck Surgery, Washington University School of Medicine; and †Department of Biomedical Engineering, Washington University School of Engineering and Applied Science, St. Louis, Missouri*

Objectives: A mixed reality (MR) headset that enables three-dimensional (3D) visualization of interactive holograms anchored to specific points in physical space was developed for use with lateral skull base anatomy. The objectives of this study are to: 1) develop an augmented reality platform using the headset for visualization of temporal bone structures, and 2) measure the accuracy of the platform as an image guidance system.

Methods: A combination of semiautomatic and manual segmentation was used to generate 3D reconstructions of soft tissue and bony anatomy of cadaver heads and temporal bones from 2D computed tomography images. A Mixed-Reality platform was developed using C# programming to generate interactive 3D holograms that could be displayed in the HoloLens headset. Accuracy of visual surface registration was determined by target registration error between seven predefined points on a 3D holographic skull and 3D printed model.

Results: Interactive 3D holograms of soft tissue, bony anatomy, and internal ear structures of cadaveric models were generated and visualized in the MR headset. Software user interface was developed to allow for user control of the virtual images through gaze, voice, and gesture commands. Visual surface point matching registration was used to align and anchor holograms to physical objects. The average target registration error of our system was 5.76 mm $\pm$ 0.54.

Conclusion: In this article, we demonstrate that an MR headset can be applied to display interactive 3D anatomic structures of the temporal bone that can be overlaid on physical models. This technology has the potential to be used as an image guidance tool during anatomic dissection and lateral skull base surgery. **Key Words:** Augmented reality—Hololens—Image-guided surgery—Mixed reality—Temporal bone—Virtual reality.

Otol Neurotol **39:**e1137–e1142, 2018.

The lateral skull base anatomy is complex and rich with essential neurovascular structures. Surgical dissection in the lateral skull base requires a firm understanding of the anatomical relationships to ensure optimal outcomes. Dissection requires careful identification of known landmarks to generate a mental image of the relative location of these structures. Safety and efficiency of dissection increases with experience as the surgeon's 3D mental map of the anatomy improves. Technology that improves visualization of anatomic structures and their spatial relationship would significantly improve the surgeon's operative experience.

Extended reality (ExR) describes the spectrum, or ''virtuality continuum'' of technology from virtual reality where the user is immersed in the digital world to augmented reality (AR) (1), which allows the user to overlay digital data or images over their environment. Recent technology advancements (2) have allowed for ExR Head Mounted Display (HMD) devices. These wearable devices make application of ExR much more accessible and applicable to everyday uses including architecture (3), healthcare (4), simulation (5), and education (6). Current headsets produced by Oculus, HTC, Sony, and Google create fully immersive virtual reality environments. In contrast, Google Glass is an example of an AR headset that can be used to display pertinent healthcare data during a medical encounter (6). Mixed

Address correspondence and reprint requests to Jonathan L. McJunkin, M.D., Department of Otolaryngology—Head and Neck Surgery, Washington University School of Medicine, St. Louis, MO 63108; E-mail: jmcjunkin@wustl.edu; Jonathan R. Silva, Ph.D., Department of Biomedical Engineering, Washington University in St. Louis, 1 Brookings Dr., CB 1097, St. Louis, MO 63108; E-mail: jonsilva@wustl.edu

Funding: Washington University in St. Louis School of Engineering and Applied Science Collaboration Initiation Grant.

Research reported in this publication was supported by the National Institute of Deafness and Other Communication Disorders within the National Institutes of Health, through the ''Development of Clinician/Researchers in Academic ENT'' training grant, award number T32DC000022.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors disclose no conflicts of interest.

DOI: 10.1097/MAO.0000000000001995

Copyright © 2018 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

reality (MR) refers to interactive digital data displayed over the native environment.

A MR HMD projects images onto a transparent lens over the user's visual field. The MR HMD has three key features ideal for image-guided navigation: 1) MR view, 2) spatial mapping, and 3) interactive hands-free interface.

Mixed Reality

The MR headset projects images onto a transparent lens overlying the user's visual field using stereoscopy to create the perception of depth for 3D models. Due to the transparency, the user maintains view of the external environment with a holographic overlay, allowing the user to maintain focus on the actual field by displaying data in a head-up manner. Current intraoperative navigation systems require the surgeon to fix their attention on a separate screen to check the relative position of a probe or instrument placed onto a point of interest. With the HMD display, the surgeon does not need to break view of the actual surgical field.

Spatial Mapping

The MR HMD employs spatial mapping technology to allow images to be anchored to specific points in the environment. The headset contains an infrared projector and camera that map the user's 3D environment when first turned on. 3D images displayed in the headset stay fixed relative to the mapped physical space while the user moves. In an operative setting, a 3D image-based model could be tethered to the patient while allowing the surgeon freedom of movement. Current intraoperative surgical navigation systems rely on coregistration of 3D data points generated from radiographic data and physical data points on the patient (7). The system then coregisters these two data sets and aligns them to allow navigation. MR allows visual registration and adjustment of the 3D holographic model as physical anatomic structures are identified.

Interactive Hands-free Interface

The MR headset features a multidimensional, hands-free interface ideal for use in a surgical setting. The MR headset contains a motion detection camera, gyrometer, and microphone. Thus, the user can control the headset by gesture, gaze, and voice commands. Application windows appear as holographs in the user's field of view similar to that of a personal computing desktop. A dot in the center of the screen acts as a cursor that can be moved around the holographic environment with head movement and gaze. Hand gestures can be used to ''click'' on objects. The headset also responds to voice commands after a voice prompt.

The objectives of this study are 1) to develop an MR platform for visualization of temporal bone structures and 2) to measure the accuracy of the platform as an image guidance system.

METHODS

The Human Research Protection Office at the study institution reviewed this study and determined that the study was exempt from Institutional Review Board review.

Image Acquisition and 3D Hologram Generation

Figure 1 demonstrates our workflow for image acquisition and 3D holographic generation. First, 0.6 mm slice thickness, high resolution, two-dimensional computed tomography (CT) scans (Siemens Biograph, Siemens Healthcare, Erlangen, Germany) of cadaveric temporal bones and heads were obtained. Digital imaging and communications in medicine (DICOM) data generated from the CT scans were imported into an open-source software ITK-SNAP (8), where a combination of semiautomatic and manual segmentation was employed to generate three-dimensional (3D) reconstruction images. Stereolithiography (STL) files of the 3D reconstructed images were then imported into an open-source system, MeshLab (9). Depending upon the complexity of the model, STL files can be comprised of thousands of raw, triangular meshes that can require substantial processing time in 3D platforms. In addition, the processing power of the headset limits the number of displayed triangular meshes. Meshlab allows for the simplification of triangular meshes while preserving geometrical detail and surface contouring. Furthermore, STL files can be converted to OBJ files in Meshlab, the latter of which is readily displayed with the MR platform described below.

Development of Mixed Reality (MR) Platform

To view and interact with 3D images on the Microsoft HoloLens (Seattle, WA), an MR platform must be built to allow for image release into the headset in the form of a Visual Studio project (Version 15.6; Microsoft, 2017). We developed a MR platform specific to our use in Unity (Version 3.1; Unity, 2017), a gaming platform for 3D animation and hologram

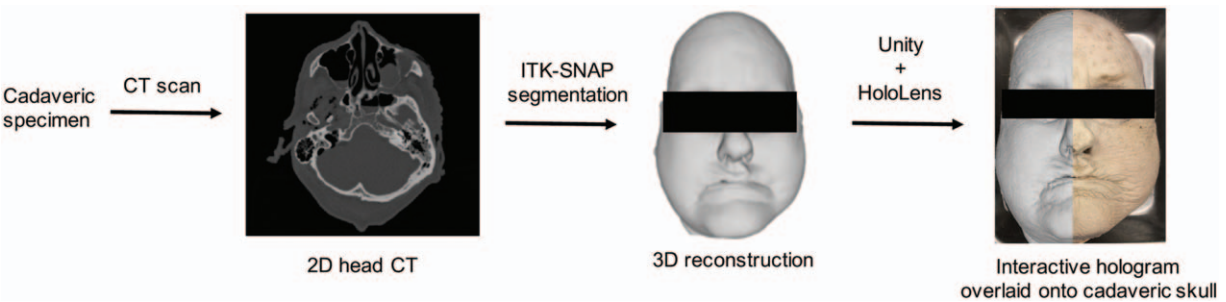

FIG. 1. Work flow demonstrating generation of 3D interactive hologram and projection onto cadaveric specimen as seen through the Microsoft HoloLens.

Copyright © 2018 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

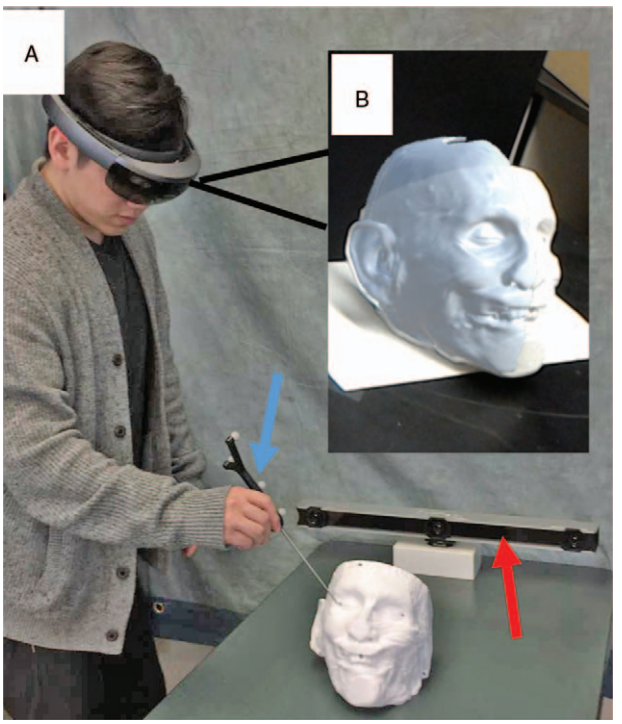

FIG. 2. Demonstration of measurement of target localization error (TLE). *A*, Observer using 4-point passive probe (blue *arrow*) to measure TLE of selected landmarks with Optitrack camera bar (red *arrow*). *B*, View through the HoloLens during measurement.

manipulation. C# scripts were used to generate interactive OBJ files that could be manipulated in the *xyz* dimensions. The images generated in Unity were then visualized in the HMD. Using C# programming in Unity and Visual Studio, code was written to allow for user-friendly voice, gaze, and gesture commands.

Surface Registration and Accuracy Measurements

Before image deployment, the HMD was calibrated each time for each specific user's interpupillary distance. Once the 3D holographic image was displayed in the HMD, soft-tissue surface anatomy was used to align the holographic images over the physical objects. A combination of voice and hand gestures was employed to manipulate the holographic images in space to overlay the physical objects.

To verify the accuracy of our system, target registration error (TRE) was measured. TRE is defined as the displacement between an image (is) and physical space (ps) location of a specific point, and is calculated using the following formula: $([X_{is} - X_{ps}]^2 + [Y_{is} - Y_{ps}]^2 + [Z_{is} - Z_{ps}]^2)^{1/2}$. We used a 3D holographic image of a cadaveric skull overlayed on a 3D printed model of the same specimen to perform our accuracy measurements. Seven landmarks, five external and two internal structures were chosen.

Our approach to error measurement is shown in Figure 2. First, a rigid four-ball passive probe was registered with an Optitrack camera bar (NaturalPoint, Coravllis, OR). Using infrared (IR) cameras, the Optitrack bar allows for tracking of marked rigid bodies. We reassigned the registration point on the passive probe from its center of mass to the tip to facilitate TRE measurement. Motive, an open-source Unity plug-in made available through Optitrack, allowed positional data of the location of the passive probe to be streamed real-time into Unity, allowing us to simultaneously record *x, y,* and *z* coordinates.

RESULTS

3D Hologram Generation and Registration for the MR HMD

Using our novel MR platform, we were able to successfully display interactive holograms of soft tissue, bony anatomy, and internal ear structures onto cadaveric and 3D printed models.

For cadaveric skull and temporal bones, the following structures were generated as interactive holograms: skin, bone, and inner ear anatomy including the facial nerve, cochlea, semicircular canals, sigmoid sinus, and internal carotid artery. We were able to successfully use voice, gesture, and gaze commands to align our holograms onto our physical objects using surface registration (Fig. 3).

C# scripts allowed the MR headset to recognize different voice, gesture, and gaze commands. Gaze commands allowed for placement of the object in two dimensions. Gesture commands were used to allow holographic images to move and rotate in all three *xyz* coordinate directions. Voice commands, e.g., ''lock,'' ''unlock'' allowed for stability of object placement in relation to its surrounding, and also disabled simultaneous manipulation by gesture or gaze commands. Our C# scripts also created a pseudo-state for each object depending on the command. For example, an image can

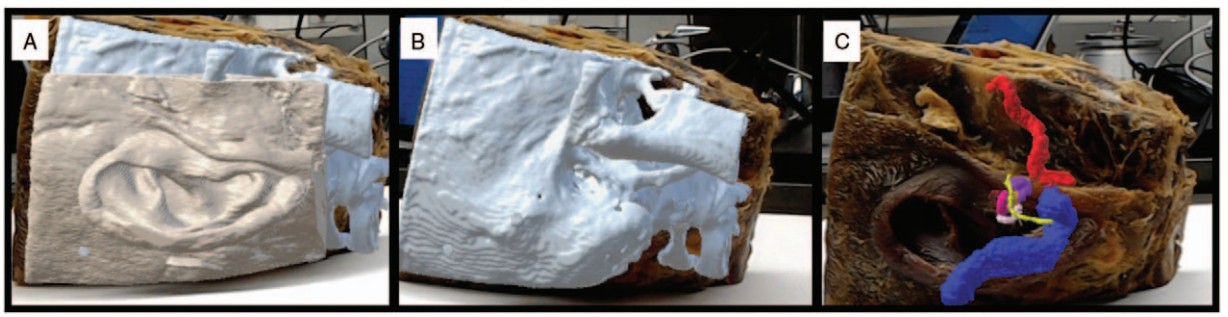

FIG. 3. 3D interactive holograms superimposed on cadaveric temporal bone as seen through the HoloLens. *A*, Skin and bone. *B*, Bone only. *C*, Internal Ear structures including jugular bulb (blue), carotid (red), facial nerve (yellow).

Copyright © 2018 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 1. *Average target localization error (TLE) of external and internal landmarks as measured by three different observers.*

Measurement Points	Average TLE (mm) ± SEM		
	Observer 1	Observer 2	Observer 3
Left earlobe	4.32 ± 0.25	5.69 ± 0.10	4.89 ± 0.11
Left lateral canthus	6.18 ± 0.36	1.65 ± 0.015	4.70 ± 0.17
Midpoint maxillary central incisors	9.23 ± 0.53	2.31 ± 0.05	5.98 ± 0.04
Right lateral canthus	1.72 ± 0.10	2.04 ± 0.03	4.19 ± 0.17
Right earlobe	8.22 ± 0.47	4.13 ± 0.11	3.89 ± 0.19
Left IAC	7.65 ± 0.44	7.50 ± 0.06	7.04 ± 0.13
Right IAC	7.87 ± 0.45	9.50 ± 34	12.3 ± 0.15
Total average	6.46 ± 0.37	4.69 ± 0.10	6.14 ± 0.14

IAC indicates internal auditory canal.

initially start off in its movable state, where manipulation of the image and overlay onto a physical object can take place. However, when the command for ''lock'' is performed, the image then shifts to its locked state. This pseudo-state allows for more refined control of the holographic image as the user has specific commands for each state.

These results were reproducible in both the cadaveric head and temporal bone models as well as the 3D-printed skull model.

Accuracy Measurements

Three different observers performed triplicate measurements of seven prespecified landmarks on the 3D holographic skull and 3D-printed skull. The left earlobe, left lateral canthus, midpoint of the maxillary central incisors, right lateral canthus, and right earlobe were chosen as external landmarks, and the right and left internal auditory canal (IAC) for internal landmarks. We first used surface registration as described above to align the hologram and 3D skull model. We were then able to use voice commands to remove the external surface of our hologram revealing just the internal landmarks to perform TRE of the right and left IAC.

For the three observers, the average TRE (±sem) was 6.46 mm ± 0.37, 4.69 mm ± 0.10, and 6.14 mm ± 0.14, respectively (Table 1). The overall average TRE between the 3 observers was 5.76 mm ± 0.54.

DISCUSSION

One of the key surgical principles taught to surgical trainees is to obtain adequate exposure. Safe surgical dissection requires knowledge of relevant anatomy. With experience, surgeons build a mental 3D map of the surgical anatomy and are able to navigate more efficiently with less exposure. An accurate surgical navigation system, using ExR, can potentially enhance a surgeon's operative experience, by enabling more efficient and safe dissection.

Image-guided surgery (IGS) using ExR, follows several basic principles: 1) target anatomy 3D model generation, 2) visual display, 3) registration and tracking, and 4) accuracy measurement.

Target anatomy 3D models are typically derived from either CT or MRI images. Modern CT scanners have accuracy to less than 0.234 mm (10). The digital imaging and communications in medicine data is segmented using one of several available software programs to generate 3D reconstructions. In our study, we used semiautomatic segmentation for skin and bone surfaces and manual segmentation for individual anatomic structures of interest. Lateral skull base anatomy is well suited for image guidance application as the relevant structures are fixed in bone. The manual segmentation process can be time-consuming (11) and efforts have been made by others to automate this process for temporal bone structures (12). ExR IGS has been described in multiple fields including neurosurgery, maxillofacial, orthopedic, cardiac, and laparoscopic surgery (13). In otolaryngology, AR has been applied in sentinel lymph node biopsy (14), endoscopic sinus and skull base surgery (15,16), and cochlear implantation (17).

Visual display of digital models in ExR has been achieved in multiple ways for IGS. A computer generated (CG) image can be superimposed on real-world imagery and displayed on a monitor, which has most often been described in laparoscopic and endoscopic surgery (18). Several groups have described using a surgical microscope with digital images projected into the eyepieces stereoscopically to allow for depth perception (19,20). The MR headset projects digital images onto a transparent lens on a HMD. The headset is wireless and allows the user freedom of motion. We were able to visualize and maneuver our digital 3D anatomic images in the headset to overlay the physical specimens (Fig. 3).

Registration and tracking in IGS are essential for reliable navigation. Registration is the process of alignment of the image space with the physical space. (7) Current commercial IGS systems use either fiducial marker points (21) or surface anatomy maps (22) for registration. Fiducials can be affixed to the skin or bone of the patient before obtaining CT or MRI. In the operating room, the physical location of the fiducials is then defined with a tracking probe. Tracking is accomplished with either optical (infrared) or electromagnetic field systems. The IGS system then aligns the image space points with the physical space points as closely as

Copyright © 2018 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

possible. In our project, we visually registered a holographic image reconstruction to its physical model using the HMD. The conceptual shift when comparing the MR headset to current IGS systems is that the registration and tracking components are contained within the headset. Our next objective was to measure the accuracy of the HMD after registration.

Accuracy in IGS is most commonly measured using TRE, which refers to the distance between a physical space point and the corresponding point as located by an IGS system. Our study found TRE to be an average of 5.76 mm using the MR headset to locate points on a 3D-printed model. Previous studies comparing accuracy using different registration methods have found bone-anchored fiducials to be the most accurate. Metzger et al. (23) reported TREs of 1.13 mm with bone-anchored fiducials, 2.03 mm with skin fiducials, 3.17 mm with bony landmarks, and 3.79 mm with a splint during maxillofacial surgery. Mascott et al. (24) found TREs of 1.7 mm using bone-anchored fiducials and 4.4 mm when using surface map registration.

There are various limitations to our research and the MR HMD. When measuring the TRE, we used a 3D-printed model in place of the original cadaveric head. In addition, to view our models in the MR headset, the number of triangular meshes in our 3D reconstructions was reduced to permit image viewing at a smooth and stable frame rate. This reduction smoothens the edges of the reconstruction and could affect model accuracy in the submillimeter range.

We found that around 30,000 triangles frame rate and spatial mapping accuracy significantly decreased, causing the model to shift as the observer moves and increasing TRE. Higher resolution reconstructions can have >200,000 surfaces, and thus are not tractable for this type of display without some refinement. As computational power is continually improving with lower power draws, newer MR headset models will have increased computing power that should help with this issue. However, given that the improvement needed is more than an order of magnitude, we will likely need to produce models that have more surfaces in areas/structures of interest while decreasing resolution elsewhere. We perceive the majority of the inaccuracy is due to image distortion and difficulty with accurate depth perception. The MR HMD creates the perception of depth by presenting images stereoscopically—a separate image to each eye on the lens. Visualizing the location of a probe in three planes is difficult without moving to different perspectives. The user can experience discomfort trying to localize points in space for extended periods. Inattentional blindness—missing abnormal events when the user's attention is focused elsewhere—is a concern with AR surgery (25). It is difficult to describe the user experience in static format as in the figures and we would encourage the reader to view this video (https://wustl.box.com/s/38o05xoygr50hmbf3ou4nj7w6lk1yhel). We attempted to capture the utility of this platform with objective measurements (TRE), but the potential application is more readily apparent with headset demonstration.

Despite the limitations identified with the MR HMD as an IGS tool, we think that this technology is promising and warrants further development. We plan to improve our surface registration method by improving the user interface to make fine adjustments of the hologram easier. Autoregistration will require defining the region of physical space where our image model will be anchored to be better mapped by the headset.

CONCLUSION

The MR HMD is capable of providing ''x-ray'' vision of human anatomy. The logistics of registration, navigation, and accuracy warrant improvement before it is ready to be used for image-guided surgery. We are currently working to improve each of these areas. The device has several key features that make it ideal for IGS. The technology has significant promise to improve surgical navigation by helping the surgical trainee more efficiently develop a mental 3D anatomical map. Future iterations of the MR headsets may well be capable of reliable registration and better accuracy for intraoperative surgical guidance to improve surgical outcomes and improve physician training.

REFERENCES

1. Milgram P, Kishino F. A taxonomy of mixed reality visual-displays. *IEICE T Inf Syst* 1994;E77d:1321–9.
2. Silva JNA, Southworth M, Raptis C, Silva J. Emerging applications of virtual reality in cardiovascular medicine. *JACC Basic Transl Sci* 2018;3:420–30.
3. Bae H, Golparvar-Fard M, White J. High-precision vision-based mobile augmented reality system for context-aware architectural, engineering, construction and facility management (AEC/FM) applications. *Visualization* 2013;1.1:3.
4. Carrino F, Rizzotti D, Gheorghe C, Bakajika PK, Fracescotti-Paquier F, Mugellini E. *Augmented reality treatment for phantom limb pain''. International Conference on Virtual, Augmented and Mixed Reality*. Springer; 2014. 248–257.
5. Fang TY, Wang PC, Liu CH, Su MC, Yeh SC. Evaluation of a haptics-based virtual reality temporal bone simulator for anatomy and surgery training. *Comput Methods Programs Biomed* 2014;13:674–81.
6. Liebert CA, Zayed MA, Aalami O, Tran J, Lau JN. Novel use of google glass for procedural wireless vital sign monitoring. *Surg Innov* 2016;23:366–73.
7. Labadie RF, Fitzpatrick JM. *Image-Guided Surgery*. San Diego, CA: Plural Publishing; 2016.
8. Yushkevich PA, Piven J, Hazlett HC, et al. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage* 2006;31:1116–28.
9. Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G. MeshLab: an open-source mesh processing tool. *Sixth Eurographics Italian Chapter Conference* 2008;129–36. www.meshlab.net.
10. Dillon NP, Siebold MA, Mitchell JE, et al. Increasing safety of a robotic system for inner ear surgery using probabilistic error modeling near vital anatomy. *Proc SPIE Int Soc Opt Eng* 2016;9786:97861G.
11. Nakashima S, Sando I, Takahashi H, Fujita S. Computer-aided 3-D reconstruction and measurement of the facial canal and facial nerve. I. cross-sectional area and diameter: Preliminary report. *Laryngoscope* 1993;103:1150–6.

Copyright © 2018 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.

12. Noble JH, Dawant BM, Warren FM, Labadie RF. Automatic identification and 3D rendering of temporal bone anatomy. *Otol Neurotol* 2009;30:436–42.
13. Kersten-Oertel M, Jannin P, Collins DL. The state of the art of visualization in mixed reality image guided surgery. *Comput Med Imaging Graph* 2013;37:98–112.
14. Profeta AC, Schilling C, McGurk M. Augmented reality visualization in head and neck surgery: An overview of recent findings in sentinel node biopsy and future perspectives. *Br J Oral Maxillofac Surg* 2016;54:694–6.
15. Li L, Yang J, Chu Y, et al. A novel augmented reality navigation system for endoscopic sinus and skull base surgery: A feasibility study. *PLoS One* 2016;11:e0146996.
16. Citardi MJ, Agbetoba A, Bigcas JL, Luong A. Augmented reality for endoscopic sinus surgery with surgical navigation: A cadaver study. *Int Forum of Allergy Rhinol* 2016;6:523–8.
17. Liu WP, Azizian M, Sorger J, et al. Cadaveric feasibility study of da Vinci Si-assisted cochlear implant with augmented visual navigation for otologic surgery. *JAMA Otolaryngol Head Neck Surg* 2014;140:208–14.
18. Vávra P, Roman J, Zonča P, et al. Recent development of augmented reality in surgery: A review. *J Healthc Eng* 2017;2017: 4574172.
19. Birkfellner W, Figl M, Matula C, et al. Computer-enhanced stereoscopic vision in a head-mounted operating binocular. *Phys Med Biol* 2003;48:N49–57.
20. Edwards PJ, Hawkes DJ, Hill DL, et al. Augmentation of reality using an operating microscope for otolaryngology and neurosurgical guidance. *J Image Guid Surg* 1995;1:172–8.
21. Grunert P, Müller-Forell W, Darabi K, et al. Basic principles and clinical applications of neuronavigation and intraoperative computed tomography. *Comput Aided Surg* 1998;3:166–73.
22. Pelizzari CA, Chen GT, Spelbring DR, Weichselbaum RR, Chen CT. Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. *J Comput Assist Tomogr* 1989;13:20–6.
23. Metzger MC, Rafii A, Holhweg-Majert B, Pham AM, Strong B. Comparison of 4 registration strategies for computer-aided maxillofacial surgery. *Otolaryngol Head Neck Surg* 2007;137:93–9.
24. Mascott CR, Sol J-C, Bousquet P, Lagarrigue J, Lazorthes Y, Lauwers-Cances V. Quantification of true in vivo (application) accuracy in cranial image-guided surgery: Influence of mode of patient registration. *Neurosurgery* 2006;59 (1 suppl 1). ONS146-56–discussion ONS146-56.
25. Dixon BJ, Daly MJ, Chan HHL, Vescan A, Witterick IJ, Irish JC. Inattentional blindness increased with augmented reality surgical navigation. *Am J Rhinol Allergy* 2014;28:433–7.

Copyright © 2018 Otology & Neurotology, Inc. Unauthorized reproduction of this article is prohibited.