



Published in final edited form as:

Int J Comput Assist Radiol Surg. 2017 June ; 12(6): 901–910. doi:10.1007/s11548-017-1564-y.

Comparison of optical see-through head-mounted displays for surgical interventions with object-anchored 2D-display

Long Qian^{1,4}, Alexander Barthel^{1,2}, Alex Johnson³, Greg Osgood³, Peter Kazanzides⁴, Nassir Navab^{1,2}, and Bernhard Fuerst¹

¹Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD, USA

²Computer Aided Medical Procedures, Technische Universität München, Munich, Germany

³Department of Orthopaedic Surgery, Johns Hopkins Hospital, Baltimore, MD, USA

⁴Laboratory for Computational Sensing and Robotics, Johns Hopkins University, Baltimore, MD, USA

Abstract

Purpose—Optical see-through head-mounted displays (OST-HMD) feature an unhindered and instantaneous view of the surgery site and can enable a mixed reality experience for surgeons during procedures. In this paper, we present a systematic approach to identify the criteria for evaluation of OST-HMD technologies for specific clinical scenarios, which benefit from using an object-anchored 2D-display visualizing medical information.

Methods—Criteria for evaluating the performance of OST-HMDs for visualization of medical information and its usage are identified and proposed. These include *text readability*, *contrast perception*, *task load*, *frame rate*, and *system lag*. We choose to compare three commercially available OST-HMDs, which are representatives of currently available head-mounted display technologies. A multi-user study and an offline experiment are conducted to evaluate their performance.

Results—Statistical analysis demonstrates that Microsoft HoloLens performs best among the three tested OST-HMDs, in terms of contrast perception, task load, and frame rate, while ODG R-7 offers similar text readability. The integration of indoor localization and fiducial tracking on the HoloLens provides significantly less system lag in a relatively motionless scenario.

Conclusions—With ever more OST-HMDs appearing on the market, the proposed criteria could be used in the evaluation of their suitability for mixed reality surgical intervention. Currently,

Correspondence to: Long Qian.

L. Qian and A. Barthel: contributed equally and should be considered joint first authors

Compliance with ethical standards

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article contains a study with human participants, which was approved by the JHU Homewood Institutional Review Board under the numbers HIRB00003228 and HIRB00004665.

Informed consent: Informed consent was obtained from all individual participants included in the study.

Microsoft HoloLens may be more suitable than ODG R-7 and Epson Moverio BT-200 for clinical usability in terms of the evaluated criteria. To the best of our knowledge, this is the first paper that presents a methodology and conducts experiments to evaluate and compare OST-HMDs for their use as object-anchored 2D-display during interventions.

Keywords

Mixed reality; Intervention; Optical see-through head-mounted display; User study

Introduction

Since the first head-mounted display (HMD)-based augmented reality system was introduced in 1968 by Ivan Sutherland [32], wearable visualization devices and mixed reality (MR) [23] technology have proven their use in military, engineering, gaming, and medical applications [27,30,31]. The advantage of HMDs as a display media over conventional monitors during intervention includes the synchronization of body, image, and action [27]. The attempts to deploy HMDs in the operating room have been under continuous investigation [28]. In this paper, a systematic approach is proposed to identify the criteria for evaluating the suitability of OST-HMDs for clinical scenarios with object-anchored 2D-display mixed reality visualization. Three OST-HMDs with different optics and system technologies are evaluated with the proposed criteria.

Classification of HMD-based mixed reality systems

To characterize different components used to build a mixed reality system for image-guided surgery (IGS), a taxonomy called data, visualization processing, view (DVV) [14,15] was proposed. In this paper, we alter the general DVV taxonomy to a more specific classification for HMD-based mixed reality interventions. Head-mounted display is differentiated into *video see-through* (VST) and *optical see-through* (OST). Instead of categorizing perception location semantically into patient, display device, surgical tool, and real environment in the DVV taxonomy, technical-oriented categories are used, such as *head-anchored*, *user-body-anchored*, *world-anchored*, and *object-anchored* based on how the virtuality is registered with the reality. The information presented on the HMD system includes *2D-display*, *3D-object*, and *perceptual augmentation*. *2D-display* refers to the display of images, like a virtual monitor placed in the environment. *3D-object* is the visualization of a model which has three dimensions. *Perceptual augmentation* refers to other kinds of virtual information, for instance spatial sound, perceptual cue of depth, or occlusion. The classification is summarized in Table 1. Based on the classification, the use cases of HMDs for mixed reality interventions mainly fall into four categories: *VST-HMD with head-anchored 2D-display* [10,16,20,24,33,41], *OST-HMD with head-anchored 2D-display* [2,3,25,29,40], *VST-HMD with object-anchored 3D-object* [1,4–6,8,21], and *OST-HMD with object-anchored 3D-object* [7,37].

Clinical use of OST-HMDs

OST-HMDs offer an unhindered and instantaneous view of the surgery scene to the surgeon, which is of critical importance in medical visualization where safety is the top concern [28].

Even if the display malfunctions, direct vision of reality would not be affected, allowing the surgeon to safely continue the operation. Recently, with the commercialization of many advanced optics designs and increasing embedded computational power, we are excited to see more consumer available OST-HMD products [18] and, accordingly, an increasing number of applications in clinical scenarios, either as *head-anchored 2D-display* [2,3,25,29,40] or *object-anchored 3D-object* [7,37]. ODG R-7¹ smart-glasses are currently being used as *head-anchored 2D-display* at the Johns Hopkins Hospital ² (Fig. 1), with early positive results in terms of surgeon comfort and satisfaction. These use cases involve streaming fluoroscopic frames from the intraoperative C-arm to the devices and displaying these images without further augmentation. Although the use cases of OST-HMDs in surgery scenarios are increasing, a frame-work to address the performance of OST-HMD devices and compare their suitability within the clinical context does not yet exist.

OST-HMD for object-anchored 2D-display

This paper focuses on one particular mixed reality surgical scenario: *OST-HMD for object-anchored 2D-display*. In the scenario of *head-anchored 2D-display*, the virtual display can be super-imposed on a critical surgery site during the surgeon's head motion. The simultaneous and clear view of both surgery site and virtual display is therefore not guaranteed. On the contrary, *object-anchored 2D-display* keeps the perception location [14] of the surgeon close to the surgery site, avoiding blocking the direct view of it. It creates a display effect similar to a virtual monitor, which resembles physical monitors that are already widely used and perceptually accepted by surgeons. Therefore, the transition from current clinical scenario to *object-anchored 2D-display* is much easier. Another mixed reality surgical scenario, *object-anchored 3D-object*, involves preoperative imaging, reconstruction, and 3D registration. Although it is potentially very helpful, it is unrealistic to bring it into routine CAI procedures using current technologies. In "Application template of object-anchored 2D-display" section, the idea of *OST-HMD for object-anchored 2D-display* is further explained, and we believe this scenario is more realistic and highly valuable for mixed reality interventions in the near future.

Contributions

In this paper, we chose three off-the-shelf OST-HMDs (Epson Moverio BT-200, ODG R-7, Microsoft HoloLens) for a combined comparison study, as they are representatives of the technologies currently available. In "Application template of object-anchored 2D-display" section, we first elaborate the model of our clinical scenario: *object-anchored 2D-display*, and continue by presenting the methodology for tracking, visualization, and calibration of the hardware used. Furthermore, the evaluation criteria, including *text readability*, *contrast perception*, *task load*, *frame rate*, and *system lag*, are introduced. In "Experiment" section, the design of the combined comparison study is described, including the multi-user study for evaluating the perceptual criteria, and an offline experiment to measure system lag in different situations. The results of the comparison study are presented and analyzed in

¹ODG R-7: <https://www.osterhoutgroup.com/products-r7-glasses>.

²Testing Smart-glasses to Superimpose Images for Surgery: <http://www.hopkinsmedicine.org/news/articles/testing-smartglasses-to-superimpose-images-for-surgery>.

“Results and discussion” section. Lastly, “Conclusion” section summarizes and concludes the paper. To the best of our knowledge, this paper is the first to assess and compare a set of OST-HMDs for mixed reality interventions with *object-anchored 2D-display*.

Methods

Application template of object-anchored 2D-display

Since the criteria of evaluating OST-HMDs is dependent on the clinical scenario, an application template is designed so that the assessment can be generalized for a certain category of mixed reality interventions. In this paper, the scenario of *object-anchored 2D-display* is chosen due to its ergonomic benefit, technical readiness, and the ease of transition from current clinical scenario. *Object-anchored 2D-display* scenario is applicable to many image-guided surgeries. Consider intramedullary nailing of the femur for example, fluoroscopic imaging is primarily needed during the placement of the starting guide wire and when distal interlocking screws are placed. Object-anchored 2D-display at the proximal and distal portions of the limb would provide optimal positioning for the critical portions of the procedure that require fluoroscopy while leaving the field of view unobstructed for those portions that do not. Knee arthroscopy is another case where object-anchored could be preferred to head-anchored. The virtual screen could be locked onto the knee and positioned in a way that provides optimal contrast with the background while maintaining proper ergonomics for the surgeon. With head-anchored display, obtaining this optimal contrast may require the surgeons head to be in an awkward or uncomfortable position.

In the preoperative stage of this category of this scenario, *display calibration* is a necessary precondition, while *processing of preoperative medical data* and *generation of other display elements* are procedure-dependent. At the intraoperative stage, necessary tasks include the *retrieval of the medical images*, the *tracking of the object* that the 2D-display is anchored to, as well as the *rendering* based on the spatial information. The design of the mixed reality intervention template for OST-HMDs with *object-anchored 2D-display* is depicted in Fig. 2.

Tracking and visualization

For the virtuality to align correctly with the user’s perception of reality, it is necessary to perform a display calibration of the OST-HMD [13] before its usage. Single point active alignment method (SPAAM) is one of the most widely performed methods due to its simplicity and accuracy [35]. Recent efforts further increase the accuracy and ergonomics of SPAAM [26], which is desirable for interventional uses.

To implement an *object-anchored 2D-display*, a stable tracking system is needed. As many HMDs already have a front-facing video camera built in, optical tracking of an image target is used in the proposed clinical scenario. No external tracking system has to be installed in the operating room, and the line of sight from the surgeon to the surgical site is not obstructed most of the time. The tracking of an image target is realized by feature-based tracking [36]. Extracting these features from a video feed in real time is essential for displaying convincing augmentations when movements of the head or the tracking target occur (Fig. 3).

Optical tracking can be aided by other tracking systems available in the setup, e.g., inertial sensing and environment tracking, in order to overcome its limitations in accuracy, update rate, and computational effort. Inertial sensors provide compensation for user motion. Simultaneous localization and mapping (SLAM) [34] algorithms with environment sensing units (e.g., laser, camera) are able to provide geometry information of the device within the world coordinate system.

Evaluation criteria

Criteria for evaluating OST-HMD devices are proposed based on the application template designed in “Application template of object-anchored 2D-display” section: *text readability*, *contrast perception*, *task load*, *frame rate*, and *system lag*. For generality, the impact of procedure-dependent issues such as OR lights is not considered in this paper.

Text readability—The patient demographics, diagnostic information, and vitals are usually displayed as plain text. Therefore, it is necessary to evaluate how well the user is able to perceive text displayed on the OST-HMD. Text readability is user-dependent and is affected by the screen resolution, screen refresh rate, and blur introduced by the optics design [12].

Contrast perception—It is important for the surgeon to be able to distinguish even slight differences in contrast in medical images in order to facilitate decision making during the intervention. Therefore, contrast perception is proposed as one of the metrics for evaluation. Similar to text readability, contrast perception is affected mainly by the optics capability and is user-dependent.

Task load—OST-HMDs may aid users during intervention but also impose extra task load. The task load for OST-HMD visualization is affected by the ergonomics of the OST-HMD, the duration of the task, and the eye fatigue caused by the display [11], etc. NASA-TLX [9] is chosen for the assessment of task load.

Frame rate—Frame rate is critical to comfortable perception and smooth use of OST-HMDs [30]. Augmentations rendered with low frame rate cause an unpleasant experience for users. Frame rate is a comprehensive measure of the hardware capability of the OST-HMDs, and can be measured by profiling the application.

System lag—For the application scenario described in “Application template of object-anchored 2D-display” section, the system lag is the combination of the time spent on tracking, rendering, and visualization. High system lag causes unpleasant experience for the user as well, especially in terms of incorrect registration between virtuality and reality. The measurement of system lag usually requires a more capable testing platform.

Experiment

A combined comparative study, involving a multi-user study for subjective criteria, and an offline experiment of system capability, is set up in order to evaluate the performance of three OST-HMDs for *object-anchored 2D-display* during interventions. This section

compares the hardware specification of the selected OST-HMDs, describes the implementation of the application, and then depicts the system setup and the procedure for the multi-user study and the offline evaluation.

OST-HMDs for comparison

Three off-the-shelf OST-HMDs are evaluated for clinical application: Epson Moverio BT-200, ODG R-7, and Microsoft HoloLens. Each device uses a different display technology, representing three categories (projector-based, LCD projector-based, and holographic waveguide) commercially available. A summary of the hardware comparison is listed in Table 2.

Epson Moverio BT-200—Epson Moverio BT-200³ was introduced to the consumer market in 2014 and gained popularity in both research and medical communities. This device features a binocular LCD projector-based optical design. It is very lightweight (88g) and affordable for non-professional users. Experiments using the BT-200 in clinics include [1,22]. This device is chosen due to its affordable price and popularity among consumers.

ODG R-7—ODG R-7 is designed and manufactured by Osterhout design group and has been available on the consumer market since 2015. This OST-HMD has binocular projector-based optics, with a higher refresh rate (80Hz) than the BT-200 (60Hz). R-7 is suitable for professional use due to its processing power. The R-7 weighs 125g. We could not identify any literature using the R-7 in a clinical study.

Microsoft HoloLens—Microsoft HoloLens⁴ has attracted a lot of attention from both academia and industry since its public appearance in 2016. The graphical rendering on the HoloLens is powered by a customized holographic processing unit (HPU) and Windows Holographic libraries. Its holographic waveguide optics contains 2.3M holographic light points, or 2.5k light points per radian. It weighs 579g.

Augmentation, system setup, and implementation details

An image simulating an orthopedic surgery scene is attached to a blue drape, serving as a tracked target. The user wearing the OST-HMD is standing at a marked position, looking down onto the simulated surgery site. The 2D-display presented on the OST-HMD is controlled by the researcher. From the software aspect, the tracking and visualization functionalities are supported by the Vuforia SDK,⁵ and Unity3D⁶ is used to create the application. For HoloLens, the display calibration is performed with its official “Calibration” application, and the tracking is further aided by the indoor localization enabled by its environmental understanding sensors. The *object-anchored 2D-display* on the OST-HMD has a physical size of 20×20cm. The system setup is illustrated in Fig. 4.

³Epson Moverio BT-200: <https://epson.com/moverio-augmented-reality>.

⁴Microsoft HoloLens: <https://www.microsoft.com/microsoft-hololens>.

⁵Vuforia SDK: <https://www.vuforia.com/>.

⁶Unity3D Game Engine: <https://unity3d.com/>.

Sample images for the evaluation of *text readability* and *contrast perception* are demonstrated in Fig. 4 as well. A short sentence with varied font size f is placed on the 1024×1024 black background. The font size is displayed in the top left corner. Each image for the evaluation of contrast perception contains four shapes with different directions. The size of the shape is 200×200 pixels, and the size of the background 1024×1024. The contrast value c of the current image is displayed in the top left corner as well. The actual grayscale value of the shape is $1 - c$.

Although the experiment is not carried out in an operating room, efforts are made to create a similar environment. The image target showing an open orthopaedics scene can be replaced with other textures for different purposes, and can be attached to patient bed, table, or even clothes of the patient. The designed tasks do not test medical expertise, but rather the OST-HMD system. If medical images and information would be used, the evaluation of the capabilities of OST-HMDs would be biased by medical knowledge and level of expertise of the subjects. Tasks that explicitly quantify image content are therefore not suitable for our evaluation.

Multi-user evaluation of OST-HMDs

Each participant performs the experiment with all three of the OST-HMDs in a random order, minimizing the learning bias. While subjectivity in the test criteria exists, e.g., eyesight, it contributes equally for each device yielding minimal bias toward a particular device. Several series of images are presented sequentially on the *object-anchored 2D-display* for evaluation of the subjective criteria proposed in “Evaluation criteria” section. The procedure of the experiment for each participant is:

- a. The participant fills out a consent form and pre-experiment survey about the previous experience with HMDs and mixed reality.
- b. The subject is shown a series of 10 short sentences on transparent background. The subject is asked to read the sentences out loud to make sure the system is working well and the user is perceiving the test images correctly.
- c. Shapes with decreasing contrast value c are displayed (see Fig. 4) to the participant. The subject has to identify the directions of the shapes. The smallest contrast value c_{min} at which the participant is still able to tell the directions of the shapes is recorded.
- d. The subject is shown a series of short sentences with decreasing font size f . The smallest font size f_{min} is recorded for which the subject is still able to correctly read the text.
- e. The subject fills out the NASA-TLX form for the OST-HMD being used.
- f. Steps (b), (c), (d), (e) are repeated for the other two OST-HMDs.

The user study was conducted with 20 participants between the ages of 22 and 46. Participants were recruited from non-medical (13) and medical (7) students. Detailed participants demographics are shown in Table 3.

Offline evaluation of OST-HMDs

An offline experiment is set up to evaluate the *system lag* of OST-HMDs. The OST-HMD under evaluation is mounted on a tripod. A large screen showing the image target is located in front of it. The slow-motion camera (GoPro Hero4,⁷ 240Hz) is placed behind the OST-HMD, capturing both the motion of the image target and the response on the OST-HMD. The system setup is shown in Fig. 4c. The system lag for each OST-HMD is measured in three experimental situations: the image target is moving on a defined path, the image target is suddenly switching position, and just the OST-HMD is moving. Each experiment is repeated 16 times. The system lag is measured by calculating the temporal difference between the change of the environment and the response of the system, via manual annotation.

Results and discussion

Results of the user study are shown in Fig. 5. A statistical analysis was performed to study the differences of the devices in the experimental setting described. Significance is achieved for p values lower than 0.05. Normal distribution of the data is not assumed; therefore, the Friedman test is performed. The test shows significant differences in the smallest readable font size $\{f_{min}\}$ ($\chi^2(2) = 27.26, p = 0.01$), the minimal distinguishable contrast value $\{c_{min}\}$ ($\chi^2(2) = 27.24, p < 0.01$), and NASA-TLX ($\chi^2(2) = 16.95, p < 0.01$) with respect to the OST-HMD being used.

Text readability

The post hoc tests are performed using the Wilcoxon signed-rank test [39]. A Bonferroni correction for multiple comparisons is applied. Comparing each of the devices to the other ones for each of the three results yields nine hypotheses. The statistical evaluation tests $n = 9$ hypotheses with a desired level of significance $p = 0.05$. The p values to test against are therefore adjusted to $p/n = 0.05/9 = 0.0056$. The resulting z -scores are then compared to the critical z -value (2.753) for the given level of significance. To reject the null hypothesis, which says that there is no difference between the two devices, the z -score has to be greater than the absolute of the critical z -value for the given level of significance. The post hoc tests show a significant improvement of text readability from BT-200 to R-7 ($Z = 2.930, p = 0.0056$), and HoloLens yields better results compared to BT-200 ($Z = 3.510, p = 0.0056$). However, there is no significant difference between HoloLens and R-7. One of the possible reasons for HoloLens and R-7 to perform better than BT-200 is the higher screen resolution and screen refresh rate.

Contrast perception

All combinations of BT-200, R-7, and HoloLens show significant differences in the minimal contrast value that is distinguishable. This value is lower for BT-200 than R-7 ($Z = -2.880, p = 0.0056$). There are also significant improvements from R-7 to HoloLens ($Z = 3.470, p = 0.0056$) and from BT-200 to HoloLens ($Z = 3.039, p = 0.0056$). HoloLens outperforms R-7 and BT-200 in providing correct perception of low-contrast images. *Task load:* There is a

⁷GoPro Hero4: <https://gopro.com/>.

significant reduction in NASA-TLX from BT-200 to HoloLens ($Z = 3.142$, $p = 0.0056$) and from R-7 to HoloLens ($Z = 2.991$, $p = 0.0056$). The difference in NASA-TLX between BT-200 and R-7 is not significant.

Although the HoloLens is heavier than the BT-200 and the R-7, its ergonomic design is more adjustable. With correct adjustment, the weight of HoloLens is not imposed on the user's nose, but distributed around the head. However, since the experiment generally lasts 10 minutes for each OST-HMD, the effect of weight is not sufficiently evaluated for time-consuming tasks by our experiment. A few participants held the BT-200 with their hands, which indicates that the ear hook design of the BT-200 may not be sufficient to securely attach the device to the user's head. The relative motion between the OST-HMD and the user's eye may invalidate the display calibration [13].

Eye fatigue is another source of task load for users in HMD-based tasks. Vergence-accommodation conflict is identified as the major cause of visual fatigue [11]. For the projector-based optics on which the R-7 and the BT-200 are constructed, the accommodation distance is fixed at the distance of the light source, while the vergence distance is about 1m away from where the tracked object is placed. On the contrary, the HoloLens, is a multiscopic display device [18], with reduced conflict between vergence and accommodation [17].

Frame rate

The information about the frame rate of the OST-HMDs is accessed via the Unity3D profiling tool. The BT-200 on average takes 0.0407s to render a frame (SD 0.0169s). The frame rate regularly dropped to less than 20 frames per second. This results in noticeable jitter when the user moves his or her head slightly. The average time spent rendering one frame on the R-7 is 0.0124s (SD 0.0034s), and the average number of frames per second is 87.6648. The frame rate of the HoloLens device is most stable by observation. Mean render time was 0.0151s (SD 0.0029s), which corresponds to 67.7081 frames per second on average. Lower frame rates occurred only occasionally. The real-time performance of each device is illustrated in Fig. 5d.

System lag

The average system lag for each device in the three experimental situations is visualized in Fig. 5f. In the experiment with the image target moving or switching locations, the system lag of BT-200 (0.163 and 0.337s) and R-7 (0.097 and 0.204s) is smaller than HoloLens (0.193 and 0.617s). However, in the experiment that investigates lag when the OST-HMD is moving rather than the image target, the system lag of HoloLens is within one single frame of the slow-motion camera. Therefore, the lag is not measurable by the experimental setup and is significantly smaller than BT-200 (0.161s) and R-7 (0.164s). Both frame rate and system lag are mainly affected by the hardware capability and the tracking modality. Each OST-HMD is equipped with different sensors and algorithms. The joint effort of indoor location and image target tracking implemented by HoloLens results in this unique behavior of the device. For surgical scenarios, which are indoor and do not involve frequent motions

of the registration target, HoloLens might be considered a more suitable OST-HMD in terms of system lag.

In addition to HMD-based solutions, naked-eye 3D overlay with half-silvered mirror has been investigated within the context of IGS [19,38,42]. Future work includes a comprehensive comparison between OST-HMDs and naked-eye systems of their suitability for MR interventions using the proposed criteria.

Conclusion

Optical see-through head-mounted displays (OST-HMD) are able to provide a mixed reality experience while keeping an unhindered and instantaneous view of reality for the wearer [28]. With the increasing availability of OST-HMDs on the consumer market and the interest from the clinical community to deploy them [30], it is necessary to propose clinically relevant criteria to evaluate the suitability of different OST-HMDs and conduct a comparison between some commercially available OST-HMD devices.

Mixed reality interventions with HMDs can be categorized by the type of HMD being used, the registration between virtuality and reality, and the type of augmentation. This paper aims at mixed reality interventions with *object-anchored 2D-display* on OST-HMDs. The model of this clinical scenario is introduced, with tracking, visualization, and display calibration methods specified. Evaluation criteria for OST-HMDs with this clinical scenario are then proposed: *text readability, contrast perception, task load, frame rate, and system lag*.

Epson Moverio BT-200, ODG R-7, and Microsoft Holo-Lens were selected to be assessed by our comparative multi-user study. These devices were chosen as they are representatives of currently available technologies. Twenty participants were recruited for the multi-user study to evaluate the perceptual performance of each OST-HMD, and an offline experiment was conducted to directly evaluate the system lag. The application setup followed the *object-anchored 2D-display* scenario. Results demonstrate that HoloLens outperforms R-7 and BT-200 in contrast perception, task load, and frame rate. For text readability, there is no significant difference between HoloLens and R-7, and they both outperform BT-200. The integration of localization and optical tracking on HoloLens yields significantly smaller system lag in the situation where the OST-HMD is moving in an indoor environment. Based on our analysis, HoloLens has better performance in the proposed scenario at present. However, the clinical benefit of OST-HMDs during a particular intervention still has to be determined by procedure-specific experiments.

To the best of our knowledge, this is the first paper that presents a methodology and conducts experiments to evaluate and compare OST-HMDs for their use as object-anchored 2D-display during interventions. There are still other aspects to consider, such as extended wearability and display calibration. However, it is our objective that the choice of HMDs for interventions is scientifically justified and the design of future HMDs for interventions follows a specific clinical scenario.

Acknowledgments

Funding This study was funded by NIAMS of the National Institutes of Health under award number T32AR067708.

References

1. Abe Y, Sato S, Kato K, Hyakumachi T, Yanagibashi Y, Ito M, Abumi K. A novel 3D guidance system using augmented reality for percutaneous vertebroplasty: technical note. *J Neurosurg Spine*. 2013; 19(4):492–501. [PubMed: 23952323]
2. Armstrong DG, Rankin TM, Giovinco NA, Mills JL, Matsuoka Y. A heads-up display for diabetic limb salvage surgery a view through the google looking glass. *J Diabetes Sci Technol*. 2014; 8:951–956. [PubMed: 24876445]
3. Azimi E, Doswell J, Kazanzides P. Augmented reality goggles with an integrated tracking system for navigation in neurosurgery. 2012 IEEE Virtual Reality Workshops (VRW). 2012; :123–124. DOI: 10.1109/VR.2012.6180913
4. Badiali G, Ferrari V, Cutolo F, Freschi C, Caramella D, Bianchi A, Marchetti C. Augmented reality as an aid in maxillofacial surgery: validation of a wearable system allowing maxillary repositioning. *J Craniomaxillofacial Surg*. 2014; 42(8):1970–1976.
5. Bajura M, Fuchs H, Ohbuchi R. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *SIGGRAPH*. 1992; 26(2):203–210. DOI: 10.1145/142920.134061
6. Bichlmeier, C., Ockert, B., Heining, SM., Ahmadi, A., Navab, N. Stepping into the operating theater: ARAV—augmented reality aided vertebroplasty. Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. ISMAR '08; Washington, DC: IEEE Computer Society; 2008. p. 165–166.
7. Chen X, Xu L, Wang Y, Wang H, Wang F, Zeng X, Wang Q, Egger J. Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display. *J Biomed Inform*. 2015; 55:124–131. [PubMed: 25882923]
8. Cutolo F, Parchi PD, Ferrari V. Video see through AR head-mounted display for medical procedures. 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). 2014; :393–396. DOI: 10.1109/ISMAR.2014.6948504
9. Hart, SG. Proceedings of the human factors and ergonomics society annual meeting. Vol. 50. Sage Publications; 2006. NASA-task load index (NASA-TLX); 20 years later; p. 904–908.
10. Herron D, Lantis J II, Maykel J, Basu C, Schwaitzberg S. The 3-D monitor and head-mounted display. *Surg Endosc*. 1999; 13(8):751–755. [PubMed: 10430678]
11. Hoffman DM, Girshick AR, Akeley K, Banks MS. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *J Vision*. 2008; 8(3):33–33.
12. Itoh Y, Amano T, Iwai D, Klinker G. Gaussian light field: estimation of viewpoint-dependent blur for optical see-through head-mounted displays. *IEEE Trans Vis Comput Graph*. 2016; 22(11): 2368–2376. [PubMed: 27479971]
13. Janin AL, Mizell DW, Caudell TP. Calibration of head-mounted displays for augmented reality applications. Proceedings of VR annual international symposium. 1993; :246–255. DOI: 10.1109/VRAIS.1993.380772
14. Kersten-Oertel M, Jannin P, Collins DL. DVV: a taxonomy for mixed reality visualization in image guided surgery. *IEEE Trans Vis Comput Graph*. 2012; 18(2):332–352. [PubMed: 21383411]
15. 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(2):98–112. [PubMed: 23490236]
16. Koesveld J, Tetteroo G, Graaf E. Use of head-mounted display in transanal endoscopic microsurgery. *Surg Endosc*. 2003; 17(6):943–946. [PubMed: 12632129]
17. Kramida G. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Trans Vis Comput Graph*. 2016; 22(7):1912–1931. [PubMed: 26336129]
18. Kress B, Starner T. A review of head-mounted displays (HMD) technologies and applications for consumer electronics. *Proc SPIE*. 2013; 8720:87200A–87200A-13. DOI: 10.1117/12.2015654

19. Liao H, Inomata T, Sakuma I, Dohi T. Three-dimensional augmented reality for mriguided surgery using integral videography auto stereoscopic-image overlay. *IEEE Trans Biomed Eng.* 2010; 57(6):1476–1486. [PubMed: 20172791]
20. Maithel S, Villegas L, Stylopoulos N, Dawson S, Jones D. Simulated laparoscopy using a head-mounted display vs traditional video monitor: an assessment of performance and muscle fatigue. *Surg Endosc.* 2005; 19(3):406–411. [PubMed: 15624063]
21. Martin-Gonzalez A, Heining SM, Navab N. Head-mounted virtual loupe with sight-based activation for surgical applications. 2009 8th IEEE international symposium on mixed and augmented reality. 2009:207–208.
22. Mentler T, Wolters C, Herczeg M. Use cases and usability challenges for head-mounted displays in healthcare. *Curr Dir Biomed Eng.* 2015; 1(1):534–537.
23. Milgram P, Colquhoun H. A taxonomy of real and virtual world display integration. *Mixed Real Merg Real Virtual World.* 1999; 1:1–26.
24. Morisawa T, Kida H, Kusumi F, Okinaga S, Ohana M. Endoscopic submucosal dissection using head-mounted display. *Gastroenterology.* 2015; 149(2):290–291. [PubMed: 26055137]
25. Ortega G, Wolff A, Baumgaertner M, Kendoff D. Usefulness of a head mounted monitor device for viewing intraoperative fluoroscopy during orthopaedic procedures. *Arch Orthop Trauma Surg.* 2008; 128(10):1123–1126. [PubMed: 18000674]
26. Qian L, Winkler A, Fuerst B, Kazanzides P, Navab N. Reduction of interaction space in single point active alignment method for optical see-through head-mounted display calibration. 2016 IEEE international symposium on mixed and augmented reality (ISMAR-Adjunct). 2016:156–157.
27. Queisner M. Medical screen operations: how head-mounted displays transform action and perception in surgical practice. *Mediat.* 2016; 6(1):30–51.
28. Rolland JP, Fuchs H. Optical versus video see-through head-mounted displays in medical visualization. *Presence Teleoperator Virtual Environ.* 2000; 9(3):287–309.
29. Sadda P, Azimi E, Jallo G, Doswell J, Kazanzides P. Surgical navigation with a head-mounted tracking system and display. *Stud Health Technol Inform.* 2012; 184:363–369.
30. Shuhaiber JH. Augmented reality in surgery. *Arch Surg.* 2004; 139(2):170–174. [PubMed: 14769575]
31. Suthau T, Vetter M, Hassenpflug P, Meinzer HP, Hellwich O. A concept work for augmented reality visualisation based on a medical application in liver surgery. *ISPRS Arch.* 2002; 34(5):274–280.
32. Sutherland, IE. Proceedings of the fall joint computer conference, part I. ACM; 1968. A head-mounted three dimensional display; p. 757-764.9–11 December 1968
33. Thomsen MN, Lang RD. An experimental comparison of 3-dimensional and 2-dimensional endoscopic systems in a model. *Arthroscopy.* 2004; 20(4):419–423. [PubMed: 15067283]
34. Thrun, S., Leonard, JJ. Simultaneous localization and mapping. In: Siciliano, B., Khatib, O., editors. Springer handbook of robotics. Springer; Berlin: 2008. p. 871-889.
35. Tuceryan, M., Navab, N. Proceedings IEEE and ACM international symposium on augmented reality. Vol. 2000. ISAR; 2000. Single point active alignment method (SPAAM) for optical see-through HMD calibration for AR; p. 149-158.
36. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D. Pose Tracking from natural features on mobile phones. Proceedings of the 7th IEEE/ACM international symposium on mixed and augmented reality; Washington, DC, USA. 2008. p. 125-134.
37. Wang H, Wang F, Leong APY, Xu L, Chen X, Wang Q. Precision insertion of percutaneous sacroiliac screws using a novel augmented reality-based navigation system: a pilot study. *Int Orthop.* 2015; 40:1–7. [PubMed: 25820838]
38. Wang J, Suenaga H, Hoshi K, Yang L, Kobayashi E, Sakuma I, Liao H. Augmented reality navigation with automatic marker-free image registration using 3-d image overlay for dental surgery. *IEEE Trans Biomed Eng.* 2014; 61(4):1295–1304. [PubMed: 24658253]
39. Wilcoxon F, Katti S, Wilcoxon RA. Critical values and probability levels for the wilcoxon rank sum test and the wilcoxon signed rank test. *Sel Table Math Stat.* 1970; 1:171–259.

40. Yoon JW, Chen RE, Han PK, Si P, Freeman WD, Pirris SM. Technical feasibility and safety of an intraoperative head-up display device during spine instrumentation. *Int J Med Robot Comput Assist Surg.* 2016; doi: 10.1002/rcs.1770
41. Yoshida S, Kihara K, Takeshita H, Nakanishi Y, Kijima T, Ishioka J, Matsuoka Y, Numao N, Saito K, Fujii Y. Head-mounted display for a personal integrated image monitoring system: ureteral stent placement. *Urol Int.* 2014; 94(1):117–120. [PubMed: 25139526]
42. Zhang X, Chen G, Liao H. High quality see-through surgical guidance system using enhanced 3d autostereoscopic augmented reality. *IEEE Trans Biomed Eng.* 2016; doi: 10.1109/TBME.2016.2624632

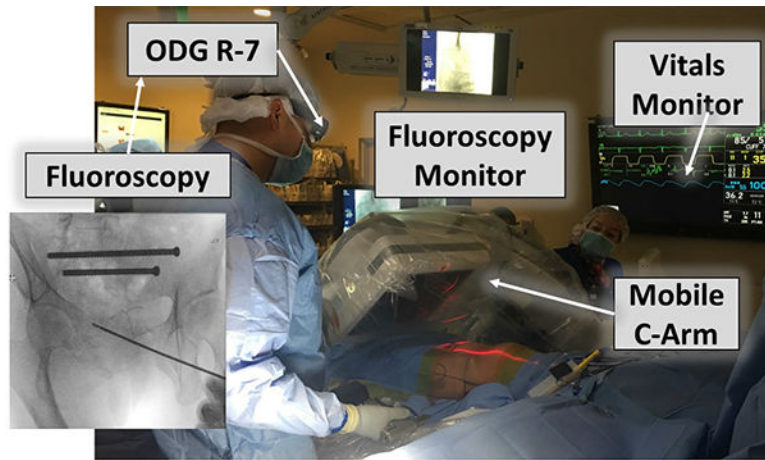


Fig. 1. Use case of ODG R-7 in orthopaedics surgery (see footnote 2) performed by our clinical partner Greg Osgood, M.D

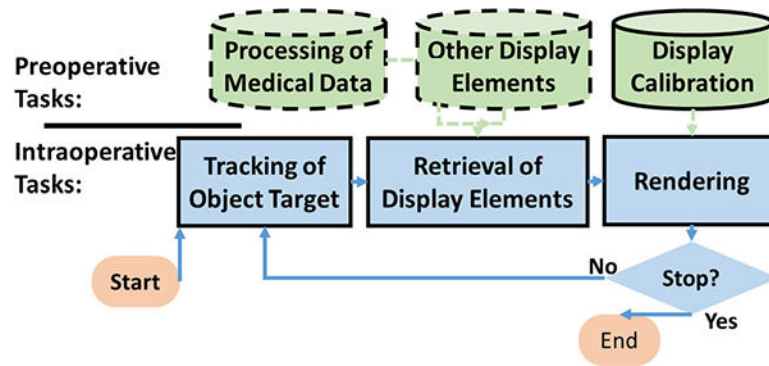


Fig. 2. Application template for OST-HMDs with object-anchored 2D-display



Fig. 3.
Three optical see-through head-mounted displays for evaluation. **a** Epson Moverio BT-200.
b ODG R-7. **c** Microsoft HoloLens

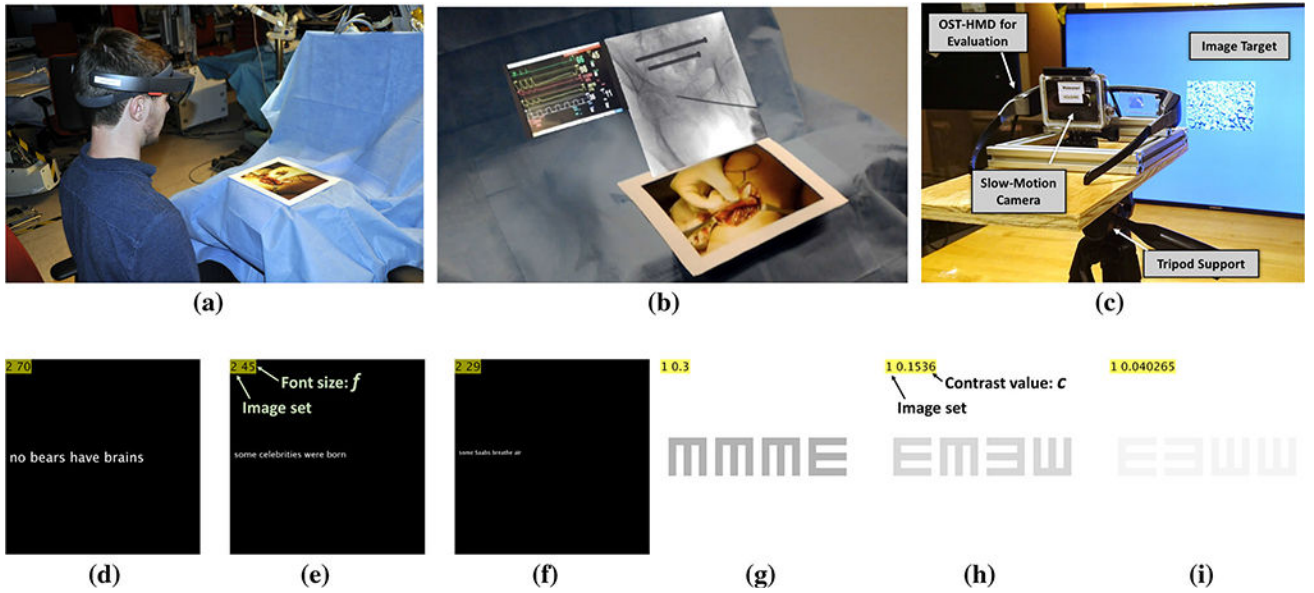


Fig. 4.
d–f Three sample images for evaluating the text readability. **g–i** Three sample images for evaluating the contrast perception. **a** Participant stands beside the simulated surgery site, testing the performance of OST-HMDs sequentially. An image is used as a tracking target. **b** Mixed reality capture of HoloLens: A 2D-display is anchored to the tracking target. The 2D-display visualizes sample images of contrast perception and text readability. **c** Experimental setup for offline evaluation of system lag. The slow-motion camera captures the motion of image target and the display on OST-HMDs

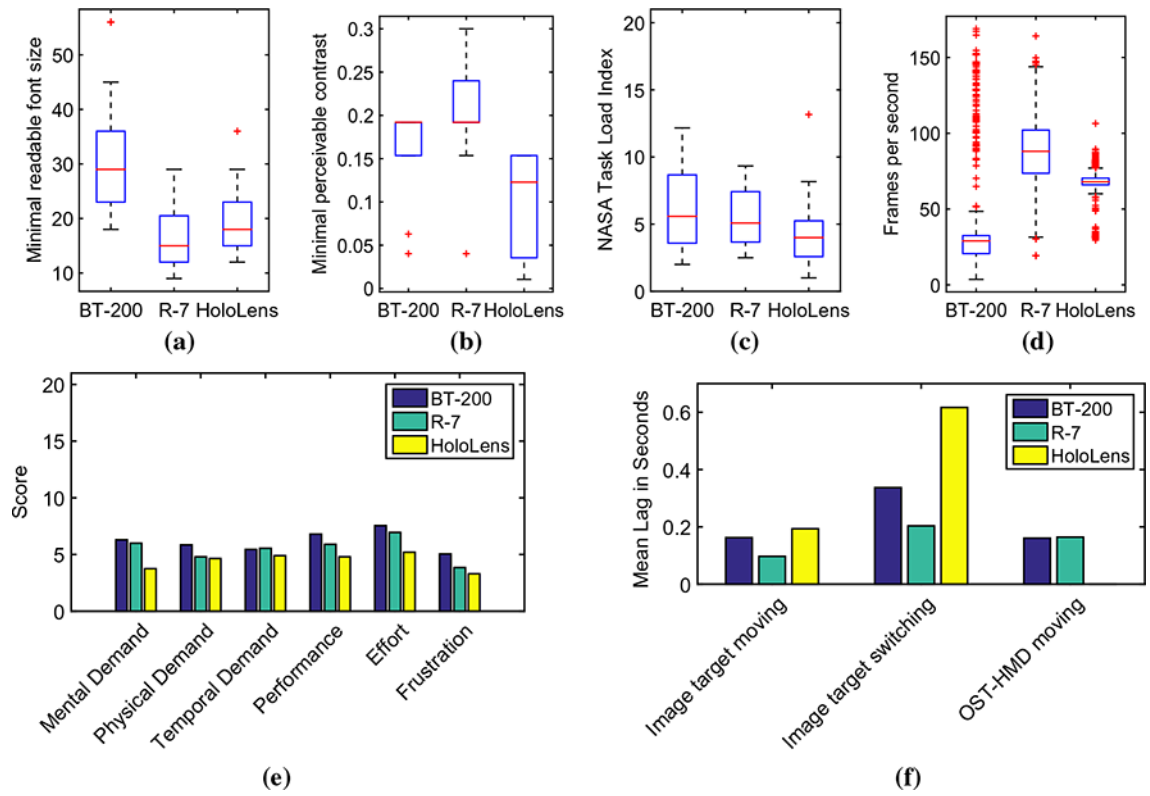


Fig. 5. Experiment results. Lower values indicate better performance. **a** Text readability, **b** contrast, **c** task load, **d** frame rate, **e** results of the NASA-TLX, **f** lag measured in the offline experiment

Table 1

Technical classification of HMD-based mixed reality interventions

HMD	Registration	Display
+ VST-HMD	+ Head-anchored	+ 2D-display
+ OST-HMD	+ User-body-anchored	+ 3D-object
	+ World-anchored	+ Perceptual augmentation
	+ Object-anchored	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Overview of selected hardware characteristics

	Moverio BT-200	ODG R-7	Microsoft HoloLens
Processor	1.2GHz dual core	2.7GHz quad core	1GHz CPU, HPU
Memory	1 GB RAM	3 GB RAM	2 GB RAM
Optical design	Projector-based with LCD	Projector-based	Holographic waveguide
Screen	Dual 960×540	Dual 1280×720	2.3M holographic light points, 2.5k/rad
Field of view	23° Diagonal	30° Diagonal	About 35°
Video resolution	640×480	1280×720	1280×720
OS	Android	ReticleOS (Android)	Windows Holographic
Weight	88g	125g	579g
Fixture	Ear hook	Overhead strap, ear hook	Overhead strap

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

Demographics of 20 participants of the user study

Gender _(males, females)	15, 5
Age _(min, max, avg, sd)	22, 46, 27, ±5.7
Handedness _(left, right, ambidextrous)	0, 19, 1
Uncorrected poor vision _(no, yes)	20, 0
Color blindness _(no, yes)	19, 1
Poor depth perception _(no, yes)	20, 0
AR experience _(none, limited, familiar, experienced)	9, 8, 2, 1
VR experience _(none, limited, familiar, experienced)	11, 9, 0, 0

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript