Integrating Eye-Tracking to Augmented Reality System for Surgical Training

Article *in* Journal of Medical Systems · September 2020

DOI: 10.1007/s10916-020-01656-w

Shang Lu
University of Alberta

Xianta Jiang
Memorial University of Newfoundland

Bin Zheng
University of Alberta

Journal of Medical Systems (2020) 44: 192
https://doi.org/10.1007/s10916-020-01656-w

EDUCATION & TRAINING

Integrating Eye-Tracking to Augmented Reality System for Surgical Training

Shang Lu[1] · Yerly Paola Sanchez Perdomo[2] · Xianta Jiang[2] · Bin Zheng[2]

Received: 15 January 2019 / Accepted: 16 September 2020 / Published online: 29 September 2020

Abstract

Augmented Reality has been utilized for surgical training. During the implementation, displaying instructional information at the right moment is critical for skill acquisition. We built a new surgical training platform combining augmented reality system (HoloLens, Microsoft) with an eye-tracker (Pupil labs, Germany). Our goal is to detect the moments of performance difficulty using the integrated eye-tracker so that the system could display instructions at the precise moment when the user is seeking instructional information during a surgical skill practice in simulation. In the paper, we describe the system design, system calibration and data transferring between these devices.

Keywords Augmented reality · Eye tracking · Calibration · Simulation · Surgical training

Introduction

Augmented Reality (AR), which adds texts, graphic images or 3D contents into the user's direct vision of the real world creating a mixed reality, has been widely used for guiding surgical procedures [2, 14, 20] and building the platforms for skill training [1–3, 22]. Many researchers [1, 2] agree that AR can provide graphic guidance to surgeons in real-time, help surgeons match images into the surgical scene, display cutting trajectories or margins, and minimize anatomical ambiguities. All these functions of the AR create substantial ways to enhance the precision of the operation for maximizing patient safety inside the operating room.

When applied to skill training, AR can guide learners to practice a complex procedure without frequent pauses in the middle of practice to search for instruction. The needed instruction can be displayed to the learners on their side vision. This feature of AR is truly helpful for healthcare learners when they are practicing a procedure composed of multiple steps, such as delivering a baby or inserting a tube to the chest for saving a life.

The crucial question in using AR for skill training is how to detect the proper moments for displaying the instructional information to maximize the learning outcome without causing a distraction. A constant display of instructional messages may take away the self-learning process of a learner. To protect the natural learning process, some systems required the learners to signal their moments of performance difficulty by giving a gesture in hand or demand by voice [4]. Both ways are not ideal as the learners need to stop the skill practice for triggering the information on the AR device. Other researchers used object tracking to detect the crucial moment during task performance, then built a context-based AR system to display a sequential instructional message at different stages of practice [5, 13, 14, 20]. However, over displaying instructional information to a learner could distract the user if the user actually does not need help. A better AR platform can detect performance difficulty of a learner based on his behaviors. Once the moment is perceived, the AR system should display needed instruction automatically. In this fashion, learners can concentrate on their practice without frequent distraction and interruption [24].

✉ Bin Zheng
bzheng1@ualberta.ca

Shang Lu
slu2@ualberta.ca

Yerly Paola Sanchez Perdomo
yerly@ualberta.ca

Xianta Jiang
xianta@ualberta.ca

1 Multimedia Research Center, Department of Computing Science, University of Alberta, Edmonton, AB, Canada

2 Surgical Simulation Research Lab, Department of Surgery, University of Alberta, Edmonton, AB, Canada

To fulfill this goal, we add an eye-tracker to the AR goggles. Eye-tracking technology provides useful pupillary and eye motion data, revealing cognitive information including human mental workload [15] and attention shift [25]. Eye tracking has been used in medical simulation research [16, 19, 21], generating important data for describing surgeons' behaviors over different stages of their expertise. Eye-tracking can also be used as an input signal for driving medical devices, such as a wheelchair and endoscope [4, 18].

Adding eye-tracker to the AR system accompanies several challenges, ranging from hardware installation, system calibration, and data processing. In this paper, we describe our work of adding the Pupil Lab Eye-tracker to the HoloLens, steps in calibrating the system for acquiring accurate eye tracking signal, and efforts in finding the moment of performance difficulty using eye-tracking data.

Implementation

In this paper, we use moments of gaze position residing on defined surfaces to trigger the display of the instructional information into the HoloLens; the pupillary information is also utilized as a potential indicator of practicing difficulty. With the help of our system, trainees can see the AR information while they are at the difficult moment during surgical training. Viewing such information at their side view, trainees are supposed to fasten their learning process.

Research environment

The whole system was built in the Surgical Simulation Research Lab (SSRL) at the University of Alberta. An AR device, an eye-tracker and a high-performance computer were connected and organized; then the system was tested on a simulation model.

HoloLens

Microsoft HoloLens is an AR device that projects visual information onto the headset screen. Normally, the HoloLens application is controlled through gesture or voice commands. We replaced these gesture interactions with gaze position, surface detection, and pupillary data. We used Unity software to build HoloLens applications and to connect the HoloLens and eye tracking software.

Pupil lab eye tracker

Pupil Lab Eye-tracker is an open source platform which is usually used to connect with Virtual Reality and AR devices [7, 12, 17, 18, 23]. Its equipment can be easily installed on HoloLens [17] with a set of add-on elements (Fig. 1). The add-on elements hold the three cameras, one for world view (world camera), and the other two for tracking eye motion (eye cameras).

The world camera has a wide angle 100-degree FOV lens, generating videos with 1280 × 720 pixels resolution and 30 frames per second. The eye cameras are No focus 200hz Eye Camera, generating 192 × 192 pixels resolution and 120 frames per second video.

Pupil lab software

Pupil Lab Eye-tracker software provides eye movement data, including gaze position and pupil diameter, as well as information of surface detected [12]. With the predefined surface, the application can easily tell whether the gaze position is located inside a certain surface.

Pupil Lab Eye-tracker software provides "HoloLens Relay" plugin which uses User Datagram Protocol (UDP) to transfer Pupil Lab Eye-tracker data to HoloLens. However, this plugin only reads and transfers gaze position information. In this study, a modification is added to this plugin to extract and transfer pupillary information.

Simulation model

We built a chest tube simulation model using a plastic mannequin to simulate a human trunk, and a soft synthetic skin pad was attached to the hard body model at the chest position (Fig. 2).

Surgical simulation task

A chest tube insertion is a lifesaving procedure and contains more than 20 steps that need to be performed in a determined sequence. Therefore, a chest tube insertion can be challenging for trainees to remember during the practice. As a result, trainees need to check the instructions during

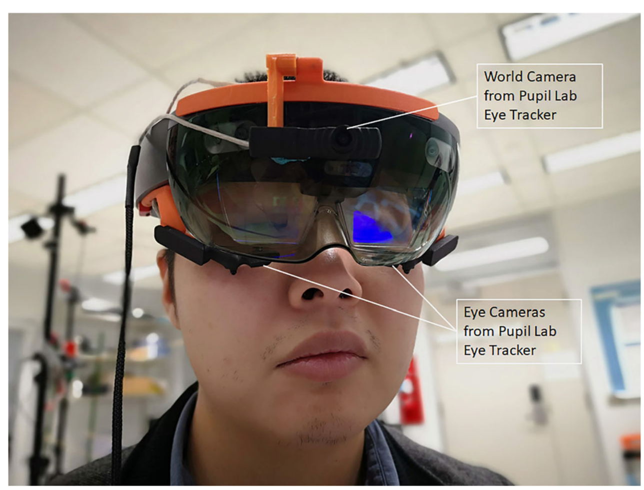


Fig. 1 HoloLens combined with Pupil Lab Eye-tracker cameras

training to make sure they are performing the task properly. Having a system that provides real-time AR instructions, trainees do not need to think hard or check the paper instructions to recall the procedure during the task. Therefore, a multistep procedure as the chest tube insertion is a suitable surgical task to test the function of our system.

Areas of interest and surface tracker

We defined several areas of interest (AOI) in the working platform. During the training task, the trainee must cut through the soft synthetic skin pad and insert the tube through the pad into the body model. In order to finish the procedure, the trainee needs to focus on the skin pad when performing. We defined the skin pad as an AOI named “Skin Pad.” Beside the working platform, we prepared the instruction sheet on a piece of cardboard. Through the HoloLens screen the trainee can look at the instruction sheet for reference when the trainee has a problem during the procedure. The instruction sheet was defined as another AOI named “Instruction.”

Pupil Lab application can help us manage the AOI easily through its official “Surface Tracker” plugin. This plugin allows users to define surfaces via markers in the environment and track them in real-time. Fig. 3 shows how the surface is prepared through markers and the detection result in the Pupil Lab application. Surfaces are registered and saved automatically into “surface definitions.” Next time Pupil Lab Eye-tracker starts, the surfaces appear on the screen like in Fig. 3 if the “Surface Tracker” plugin is on. As shown in Fig. 3, even some markers are not detected due to obscurity, the surfaces are still detected, and square windows are shown upon the surfaces. Surface data and the gaze position relative to it are pushed into data flow in Pupil Lab application. Our system can subscribe to such data and broadcast them to HoloLens (see section below).

We took the gaze position relative to surfaces as the indicator of trainee’s performance status. When gaze is located inside the skin pad, it means the trainee is probably focusing on the task on hand. Then, the AR information is disabled to eliminate distractions. When gaze is located inside the instruction sheet, it means the trainee is looking for help right now. Therefore, it is the right moment to display AR information.

System setup procedure

Eye-tracking signal acquiring and processing

To transfer pupillary and gaze position data from Pupil Lab Eye-tracker to HoloLens, we extracted the data needed from the Pupil Lab Eye-tracker data flow. Pupil Lab application deploys ZeroMQ network library [6] as the messaging bus, which is called Inter-Process Communication (IPC) backbone. One single backbone connects to multiple processes, so a one-to-many communication proxy is built (Fig. 4). Then, each process can push the message into the backbone and take others’ messages from the backbone. For example, surface data and the gaze position relative to it are generated every frame and pushed into the IPC backbone under the “surface” topic.

In the plugin we modified, an “IPC sub” socket is built, through which our plugin subscribes to the data from IPC backbone. Several topics are added into the socket. As a result, every time the data under these topics appear in the IPC backbone, our plugin catches these data.

In our plugin, a remote socket is built with UDP protocol for broadcasting data to HoloLens (Unity editor). After subscribing to data from IPC backbone, our plugin repacks the data into byte format and adds an extra header to the data indicating the type of data. Then, the plugin sends such data into the remote socket, broadcasting them to HoloLens (Unity editor).

Since Pupil Lab Eye-tracker is an open-source application, user-customized plugins can be used directly just being put into the “pupil capture settings/plugins” folder. Then the plugin can be turned on through the Pupil Lab application’s main menu.

Work flow in HoloLens

Pupil Lab Eye-tracker software officially provides the Unity package for eye tracking in HoloLens application. However, since the Pupil Lab Eye-tracker plugin was modified in this

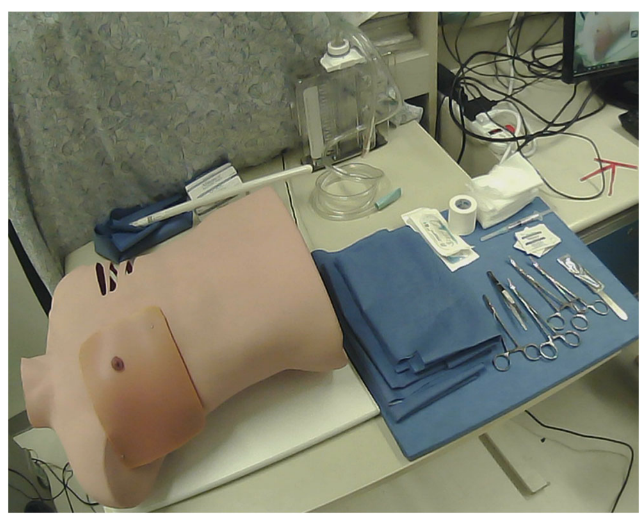


Fig. 2 Simulation setting for chest tube insertion

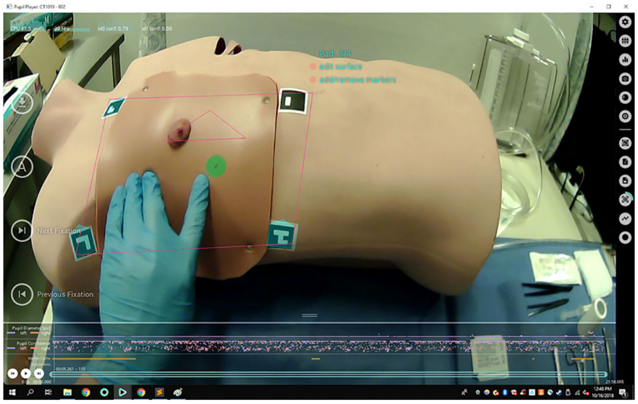


Fig. 3 Binary markers surrounding the surgical site define an area of interest (AOI)

study, a modification was needed in the HoloLens Unity package in correspondence with the new data format.

In HoloLens application (Unity editor), a UDP client socket is built, then it listens to a remote IP address which is set as the IP of Pupil Lab Eye-tracker. When HoloLens application (Unity editor) runs, it sends a byte "I3" to Pupil Lab Eye-tracker's IP address. After receiving this byte, Pupil Lab Eye-tracker responds with a byte "0I" into the broadcasting socket. When this byte is received by HoloLens (Unity editor) UDP client socket, the connection is established (Fig. 4).

In our system, all calibration and recording processes are done through Pupil Lab Eye-tracker, so we remove or skip related components in the HoloLens Unity package. After the connection is established, HoloLens application (Unity editor) sends byte "S" to eye-tracker's IP address, indicating ready for data transferring. After responding with byte "0S", when the data under selected topics are pushed into IPC backbone, Pupil Lab Eye-tracker application broadcasts all these data into the remote socket. HoloLens application (Unity editor) receives all such data through UDP client. A function is built to interpret the data from byte format to digit or string. According to the data header, the digit or string is stored under different variables and is updated every time new data is received.

Calibration

In order to extract the location of the subjects' gaze, the Pupil Lab system generates a mapping function from the pupil center detected in the eye images to gaze coordinates in the world image. This process is called calibration and should be done every time before the system is used.

First, the angles of the eye and world cameras need to be adjusted to track the whole eye area and most of the trainee's view. After angle adjustment, the trainee sits in front of the computer screen. Then, the calibration process in Pupil Lab Eye-tracker is started, using the "Screen Marker Calibration" option. A circle marker appears on the screen first at the center, then at the four corners. The trainee follows the marker with the gaze while keeping the head as still as possible. The calibration is repeated until the accuracy reaches 2.5 degrees in 3d mode.

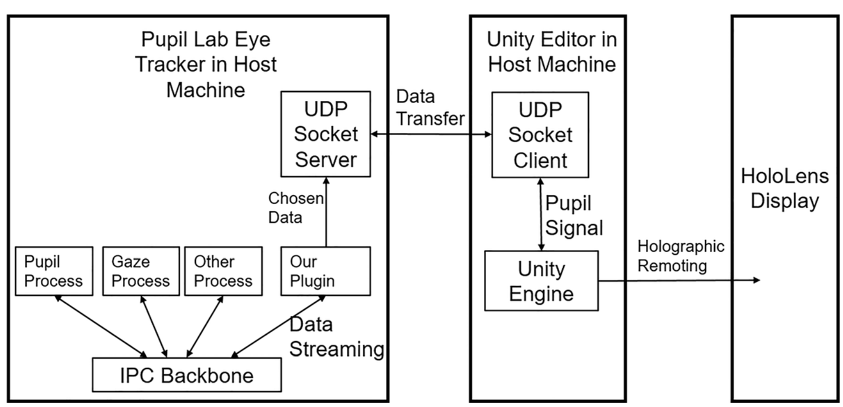
Connection setup

Pupil Lab Eye-tracker cameras are connected to the host computer through USB cable. When "Pupil Capture" is running, and our plugin is selected, the IP address of eye-tracker appears on the menu of the plugin.

The HoloLens application runs in Unity editor. In Unity, the whole project should be set on "Universal Windows Platform," with Project Setting supporting "Windows Mixed Reality" and HoloLens' "InternetClient," "InternetClientServer," "PrivateNetworkClientServer," "Microphone," and "SpatialPerception" capabilities.

HoloLens and the host computer must be connected to the same Wi-Fi network. Through the Unity asset "PupilSettings" we built, the "Pupil Remote IP" is set as the IP address of eye-tracker. Using "Holographic Remoting" in HoloLens and "Holographic Emulation Mode" in Unity, the HoloLens is connected to the host computer. After this setup, the three devices, HoloLens,

Fig. 4 Data flow between the eye-tracker and HoloLens



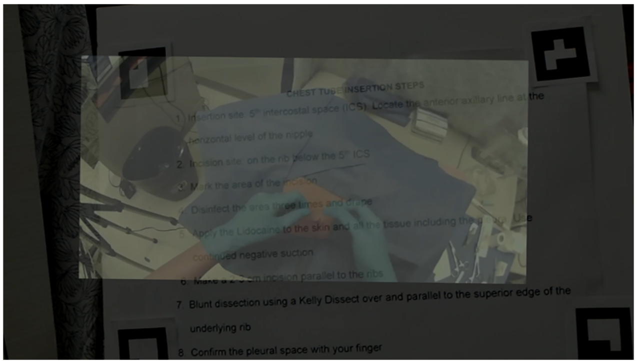


Fig. 5 Display capture of holographic animation

host computer, and Pupil Lab Eye-tracker are connected (Fig. 4).

Control of displaying AR information

After the connection setup, the Unity scene can run on the host machine in Unity editor Play mode, while the scene is broadcasted onto HoloLens through "Holographic Remoting." Now, the trainee can wear the device (HoloLens with Pupil Lab Eye-tracker cameras installed) and perform the simulation task on the model. We observed in a pilot study that trainees focused on the surgical site most of the time during the task. If a trainee had a difficulty to perform the next step or got confused about the procedure, the trainee moved their gaze from the surgical site to the instruction sheet which was set aside the simulation platform.

We defined two AOIs, one on the Skin Pad (surgical site) and the other one on the Instruction Sheet. Our system detected the position of the trainee's gaze. If the gaze position was located inside the "Instruction sheet," holographic information (texts, images, or video) appeared on the screen of HoloLens to provide instructional information to the trainee. When the gaze turned back to the surgical site the holographic information on HoloLens disappeared.

We offered two options to display the AR instructions: (1) Animation for the whole task; (2) step-by-step Images chosen by the trainee.

(1)Animation for the whole task: every time the gaze of the trainee located inside the "Instruction sheet," a short animation was displayed on the screen of Hololens (Fig. 5). Trainee can look through the animation and still see the real world, thus the animation is displayed through augmented reality. The animation consists of 26 images, each of which shows one step and lasts for 1 s. As a result, the animation lasts 26 s in total covering the whole procedures for chest tube insertion. The animation disappeared when it was over or the gaze of the trainee turned back to the surgical site.

(2) step-by-step Images chosen by the trainee: We broke the procedures of chest tube insertion into four sections: landmarks, dissection, insertion and securing. We created a holographic main menu consisting of four buttons mapping the four sections. Every time the gaze of the trainee located inside the "Instruction sheet," this menu was displayed on the screen of Hololens (Fig. 6 left). This main menu was fixed relative to the trainee's view in AR. The button changed color when the trainee's gaze was located on it and turned back to white when the gaze moved out. After the gaze stayed on a button for 5 s, the main menu was replaced by a detailed instruction panel for the corresponding section. The panel consists of several images showing every step in the section (Fig. 6 right). The same as the animation display, the main menu and the instruction panel is displayed in augmented reality. The main menu or instruction panel disappeared after displaying for 15 s or the gaze of the trainee turned back to the surgical site.

Discussion and future works

The goal for this paper is to describe our effort for tracking trainees' eye motion in the AR environment. In our project, we chose the Pupil Lab Eye-tracker and HoloLens. Pupil Lab eye-tracker is open-source and user-friendly. Since HoloLens

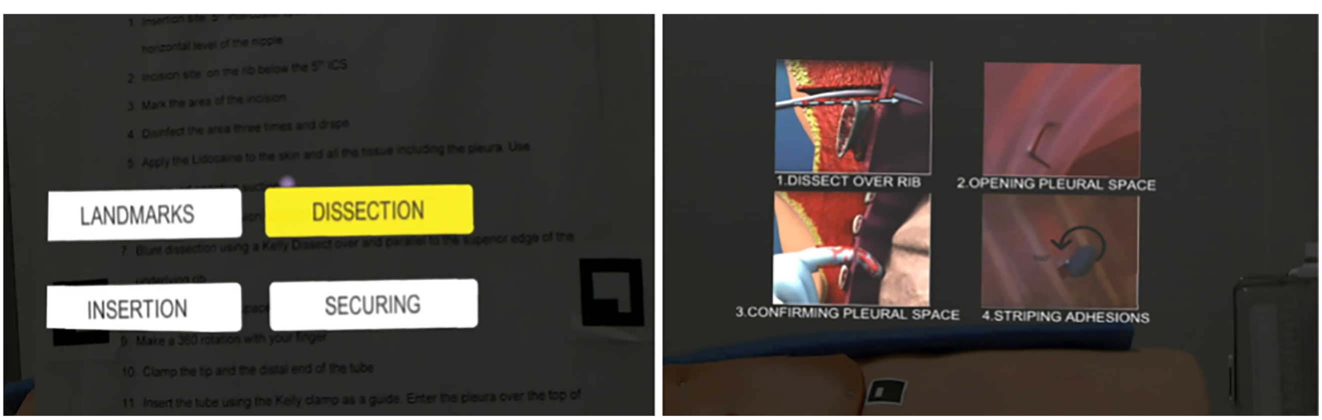


Fig. 6 Display capture of holographic step-by-step Images. Left: main menu consisting of 4 buttons. Right: instruction images for dissection as an example

is a popular AR device in the market, the system has these benefits: 1) it has a reasonable price; 2) it is convenient to work with; 3) we can run different training modules in HoloLens to train different procedures.

In our next paper, we plan to report how to process eye-tracking data to capture the moment of performance difficulty from trainees. Among eye-tracking data, pupil diameter is a source of data that can describe a user's cognitive activities. When trainees encounter a moment of practicing difficulty, we expect that their pupil may enlarge. Our team has experience in detecting the pupil dilation in surgical simulation tasks [8–11, 26]. Therefore, we plan to examine the feasibility of using pupil dilation as the trigger signal in the future. We will also examine the eye motion trajectories. When a trainee's gaze moves away from the surgical site to the instructional sheet, we may capture this moment.

Although this system provides good AR assistance on surgical training, there are some limitations. Firstly, the system depends on a UDP connection between the host computer and HoloLens. Such UDP connection can be blocked if the computer and HoloLens are connecting to different routers. As a result, the system works quite well under one router Wi-Fi, while it loses the connection under a multi-router Wi-Fi. Secondly, wearing HoloLens for a long period can produce tiredness to the trainee's eyes and discomfort due to its weight (over 500 g).

Consequently, the training session should be shortened, or rest periods need to be provided to decrease the discomforts caused by wearing the goggles. Thirdly, the purpose of this system is to help the trainee to perform the procedure without stop so that we can facilitate the learning process. However, some trainees may develop a certain level of dependency on those superimposed instructions when practicing with the AR system. We need to be aware of this habit and find a way to prevent it from happening.

Currently our system will display the whole procedure to the trainees when they look for help in the training. We believe a more advanced system which only displays the next step of the procedure will be generated in the future when artificial intelligence is involved in medical training via AR.

Acknowledgements We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) to support the technological development of this project through Discovery Grant to Dr. Zheng. We also thank the University of Alberta Provost Office for supporting the education application of this project through the Teaching & Learning Enhancement Fund to Dr. Zheng.

Compliance with ethical standards

Disclosure of potential conflicts of interest The authors declared that we do not have any conflict of interest for this research work.

Ethical approval This article does not contain any studies with animals performed by any of the authors.

References

1. Barsom E, Graafland M, Schijven M (2016) Systematic review on the effectiveness of augmented reality applications in medical training. Surgical endoscopy 30(10):4174–4183
2. Bernhardt S, Nicolau SA, Soler L, Doignon C (2017) The status of augmented reality in laparoscopic surgery as of 2016. Medical image analysis 37:66–90
3. Dickey RM, Srikishen N, Lipshultz LI, Spiess PE, Carrion RE, Hakky TS (2016) Augmented reality assisted surgery: a urologic training tool. Asian journal of andrology 18(5):732
4. Eid MA, Giakoumidis N, El-Saddik A (2016) A novel eye-gaze-controlled wheelchair system for navigating unknown environments: Case study with a person with als. IEEE Access 4:558–573
5. Esposito M, Busam B, Hennersperger C, Rackerseder J, Lu A, Navab N,Frisch B Cooperative robotic gamma imaging: Enhancing us-guided needle biopsy. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 611–618 (2015)
6. Hintjens P Ømq-the guide. Online: http://zguide.zeromq.org/page:all, Accessed on 23 (2011)
7. Jang C, Bang K, Moon S, Kim J, Lee S, Lee B (2017) Retinal 3d: augmented reality near-eye display via pupil-tracked light field projection on retina. ACM Transactions on Graphics (TOG) 36(6):190
8. Jiang X, Zheng B, Tien G, Atkins MS Pupil response to precision in surgical task execution. In: MMVR, pp. 210–214 (2013)
9. Jiang X, Atkins MS, Tien G, Bednarik R, Zheng B Pupil responses during discrete goal-directed movements. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp 2075– 2084 (2014)
10. Jiang X, Atkins MS, Tien G, Zheng B, Bednarik R Pupil dilations during target-pointing respect fitts' law. In: Proceedings of the Symposium on Eye Tracking Research and Applications, ACM, pp. 175–182 (2014)
11. Jiang X, Zheng B, Bednarik R, Atkins MS (2015) Pupil responses to continuous aiming movements. International Journal of Human-Computer Studies 83:1–11
12. Kassner M, Patera W, Bulling A Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication, ACM, pp 1151–1160 (2014)
13. Katić D, Sudra G, Speidel S, Castrillon-Oberndorfer G, Eggers G, Dillmann R Knowledge-based situation interpretation for context aware augmented reality in dental implant surgery. In: International Workshop on Medical Imaging and Virtual Reality, Springer, pp. 531–540 (2010)
14. Katić D, Wekerle AL, Görtler J, Spengler P, Bodenstedt S, Röhl S, Suwelack S, Kenngott HG, Wagner M, Müller-Stich BP, et al. (2013) Context-aware augmented reality in laparoscopic surgery. Computerized Medical Imaging and Graphics 37(2):174–182
15. Kosch T, Hassib M, Buschek D, Schmidt A Look into my eyes: Using pupil dilation to estimate mental workload for task complexity adaptation. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, p LBW617 (2018)
16. Kottayil NK, Bogdanova R, Cheng I, Zheng B, Basu A Investigation of gaze patterns in multi view laparoscopic surgery. In: Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the, IEEE, pp 4031–4034 (2016)

17. Kun AL, Meulen Hvd, Janssen CP Calling while driving: An initial experiment with HoloLens (2017)
18. Langbehn E, Steinicke F, Lappe M, Welch GF, Bruder G (2018) In the blink of an eye: leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality. ACM Transactions on Graphics (TOG) 37(4):66
19. Law BHY, Cheung PY, Wagner M, van Os S, Zheng B, Schmölzer G (2018) Analysis of neonatal resuscitation using eye tracking: a pilot study. Archives of Disease in Childhood-Fetal and Neonatal Edition 103(1):F82– F84
20. Navab N, Traub J, Sielhorst T, Feuerstein M, Bichlmeier C (2007) Action and workflow-driven augmented reality for computer-aided medical procedures. IEEE Computer Graphics and Applications 27(5):10–14
21. Patel R, Zurca A (2018) 396: I see what you see adding eye-tracking to medical simulation. Critical Care Medicine 46(1):181
22. Pelargos PE, Nagasawa DT, Lagman C, Tenn S, Demos JV, Lee SJ, Bui TT, Barnette NE, Bhatt NS, Ung N, et al. (2017) Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. Journal of Clinical Neuroscience 35:1–4
23. Renner P, Pfeiffer T Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In: 3D User Interfaces (3DUI), 2017 IEEE Symposium on, IEEE, pp 186–194 (2017)
24. Sielhorst T, Feuerstein M, Traub J, Kutter O, Navab N Campar: A software framework guaranteeing quality for medical augmented reality. International Journal of Computer Assisted Radiology and Surgery 1:29 (2006)
25. Steil J, Müller P, Sugano Y, Bulling A Forecasting user attention during everyday mobile interactions using device-integrated and wearable sensors. arXiv preprint arXiv:180106011 (2018)
26. Zheng B, Jiang X, Atkins MS (2015) Detection of changes in surgical difficulty: evidence from pupil responses. Surgical innovation 22(6):629– 635