Mixed Reality applications for the collaborative operating room – a prototypical study

N. Marahrens[1], D. Ostler[1], J. Weinzierl[1], N. Kohn[1], T. Vogel[1,2], D. Wilhelm[1,2], S. Koller[1], H. Feußner[1,2]

[1]Research Group MITI at the University Hospital rechts der Isar of the TU Munich, Trogerstr. 26, Munich, Germany
[2]Clinic and Polyclinic for Surgery at the University Hospital rechts der Isar of the TU Munich, Ismaninger Str. 22, Munich, Germany
Contact: nils.marahrens@tum.de

Introduction

The rapid development of medical technology enables new treatment methods and leads to a strong increase in minimally invasive surgical interventions. For this reason, more and more specific devices are constantly being added to the OR. At the same time the OR staff faces the challenge to manage the adjustment and timely provision of these devices during surgery [1]. Similarly, nursing teams in Germany are increasingly understaffed and are consequently filled with staff from other European countries that may have difficulties knowing the specific terms in the OR or understanding the requirements for each situation [2].

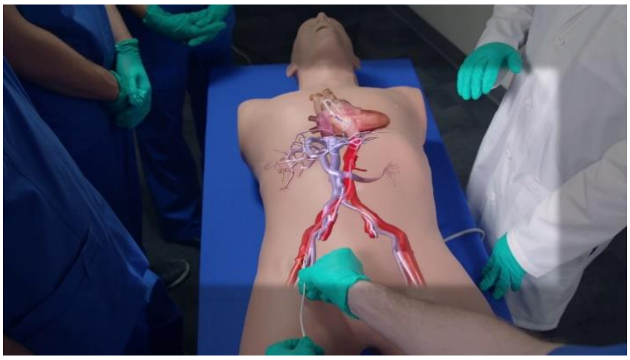

Figure 1: Exemplary application of the HoloLens for medical training (Source: caehealthcare.com)

Strategies and measures must therefore be developed to ensure and further improve efficiency, safety and outcome of the intervention. This requires solutions that bring staff and technology closer together in order to simplify system management of the OR, reduce communication issues and improve staff collaboration.

Mixed Reality (MR) applications may be a key technology to make this vision a reality. They are capable of providing information by more general means than pure language. At the same time, recent developments in machine learning allow enhanced image detection [3], workflow recognition [4] and natural language processing [5] in real time during surgery. In their combination these techniques may enable technical systems to robustly interpret the current surgical context, automatically provide the required information at the right time and react flexibly to changing demands.

In addition to providing solely information, MR application would further allow to sterilely adjust most devices in the OR from a central entity. Previously, this technology has already been used for medical training [6] (also see Fig. 1) and specific surgical applications [7]. However, most developments have mainly focused on the surgeons and less on the remaining OR staff so far.

Under these assumptions this paper investigates the potential of mixed reality applications for sterile interaction within the collaborative operating room. An exemplary use case of an assistive device that supports the scrub nurse in selecting the correct instruments is presented and evaluated in an initial user experience experiment.

Figure 2: Image of the HoloLens (Microsoft, Redmond, WA, USA) as used in the presented mixed reality application (Source: microsoft.com)

Materials and Methods

To realize a simple prototype of the envisioned system, we employed a Microsoft HoloLens (see Fig. 2). The HoloLens is a wearable device that enables the augmentation of the physical world with virtual entities. It is equipped with a multitude of sensors that enable the device to localize itself within the environment (spatial mapping) and enable various modes for user interaction.

Spatial mapping is performed by means of 4 greyscale cameras and a time-of-flight (TOF) sensor that enhances the camera images with further depth information. An additional inertial measurement unit (IMU) may also measure acceleration and spatial orientation. By combining all different sensory data, the HoloLens is able to locate itself and project holograms at any desired location within the room. All necessary sensor fusion is directly processed on the HoloLens' proprietary computer.

Aside from the possibility to interact with the system via head movements and hand gestures, the HoloLens also include four microphones to allow voice control and binaural speaker for audio feedback.

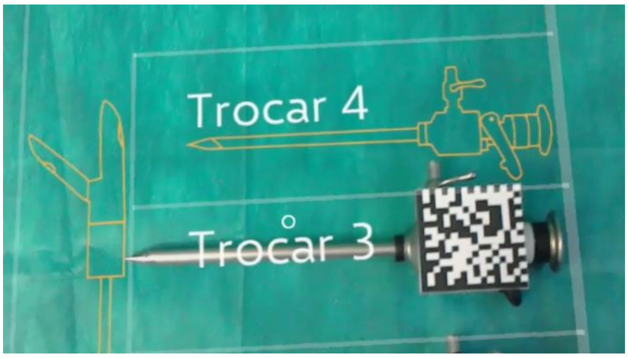

Figure 3: Segment of the grid that is projected onto the surgical tray

Included Functionalities

Our proposed assistive system supports the scrub nurse in identifying the instruments that are currently needed. It projects a grid onto the surgical tray (see Fig. 3). Each grid cell is reserved for one specific instrument and denoted by the instrument's name and contour sketch. With the integrated spatial mapping function, the grid cell pattern is automatically projected onto the surgical tray at a specified location. Through gestures the user is able to flexibly relocate the grid as desired within the plane of the surgical tray. Gestures are robustly recognized by the system even when the hands are covered with surgical gloves.

If an instruments has been placed in its corresponding grid cell, the contour sketch automatically disappears to give the user direct feedback about his or her action. The user may also interact with the system though verbal commands. By calling the name of an instrument the system clearly highlights the respective grid cell with a red box.

Currently, the integrated computer vision software of the HoloLens is not able to recognize and distinguish between the different surgical instruments in the scene. Particular problems arose due to the delicate and similar structure of some instruments. Furthermore, specular reflections from metallic surfaces highly reduced recognition accuracy. An integration of custom image processing algorithms that make use of the spatial information and could feed back their results to the HoloLens in real time are currently and to the authors' best knowledge not applicable within the API for the HoloLens. However, this functionality is crucial for the system to determine if an instrument has been correctly placed in its corresponding grid cell. For the first prototype we therefore attached QR-code markers to each instrument that were recognized within the HoloLens' video frames.

Experimental evaluation

Based on our prototype we performed a user experiment with 14 participants, some with prior OR experience. None had direct experience as a scrub nurse. During the test, participants first received a short introduction to the interaction modalities of the HoloLens such as gesture and voice control. Once participants were acquainted with the HoloLens in general, they were asked to set up the system and test out the assistive functions.

First off, users were presented the projected grid and were asked to relocate the grid to a desired position by means of gesture control. After the grid was in the right position, the participants had to place a total of 13 different surgical instruments in their respective grid cells as it is shown in Fig. 3. If the user had correctly placed the instrument in a grid cell, the contour sketch disappeared. If it did not disappear, the user had to pick up the instrument again, find the correct box and place the instrument there. Additionally, the participants were asked to interact with the system via voice commands in order to highlight the grid cells and contour sketches of specific instruments.

After successful completion the participants were all asked to fill out a User Experience Questionnaire (UEQ) [8], including a total of 26 statements that had to be rated with a 7-step Likert-scale. All statements fell into six different categories: "Attractiveness", "Perspicuity", "Efficiency", "Dependability", "Stimulation" and "Novelty".

To facilitate evaluation, we mapped the 7-step Likert-scale that takes values from 1 to 7, to values between -3 and 3 as it is also done in [9]. The resulting scale made it easier to directly tell if the user experience was rather positive (positive values) or rather negative (negative values).

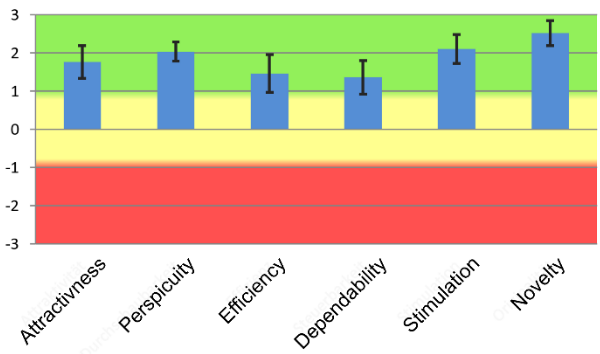

Figure 4: Results from the User Experience Questionnaire (UEQ) [8]

Aside from the user experiments voice recognition accuracy was tested for all 13 instruments by calling their respective name various times (n=30) and recording if the system recognized the command correctly.

Results

All participants were able to accomplish the task without major difficulties and therefore also completed the standardized questionnaire subsequent to the experiment. The detection of the QR code worked stable, was, however, sensitive to the distance and angle to the HoloLens' camera.

The system received mean ratings between 1.357 and 2.518. Fig. 4 shows the results as means and standard deviations over all participants and for all six categories. The scale is divided into three sections: -3 to -0.8 (very poor / poor), -0.8 to +0.8 (mediocre) and +0.8 to +3 (very good / good), marked by the respective colors red, yellow and green. Therefore, the ratings for all categories were considered as at least "good". The standard deviation takes values of less than 0.5 points in all categories. Specifically "Perspicuity", "Stimulation" and "Novelty", received the highest grades, where "Perspicuity" means how intuitive and clear the system is. Ratings for "Efficiency" and "Dependability" were the lowest, however, still fall into the highest of the three categories and were thus also considered as "good".

A relative comparison with benchmark results for typical UEQ ratings showed that for the categories "Attractiveness", "Perspicuity", "Stimulation" and "Novelty", the presented system ranks among the top 10% of the results contained in the benchmark [9]. The remaining categories "Efficiency" and "Dependability" still rank above average.

Voice detection also worked well for most instruments. Detection rates were above 70% for all instruments. All except the two keyword "PE Forceps" and "Scissors" were even detected with an accuracy of 90% or more.

Discussion

In this context, the results underline that the system was well received by most participants and its user experiences ranks well among other applications. The test clearly showed that the general concept of the system works well and that the interaction between the user and the system is intuitive. Since the test did not primarily include professional scrub nurses, this conclusion may only apply to the general usage of the system in an OR setup. Clear statements about the integration into the actual workflow and its further interaction with the additional OR team members cannot be made at this point.

Conclusions

Currently, the presented system could already be used for training, whereas a sterile application in the OR is hindered by the QR-code tags that are still necessary for instrument recognition. With further advancements in computer vision, however, a direct detection from the image may already be feasible, given that the Hololens' API allows a practical integration of custom algorithms.

Although the HoloLens is already employed for real medical applications, its weight of 579g and its relatively short battery life of 2 to 3 hours of active use might restrict their application during real surgery. Particularly for longer surgical interventions, the system might not be practical at its current state.

Future research will be devoted at its integration into the OR and the preexisting workflow. The system is planned to be coupled to a workflow recognition engine that aids at automatically highlighting the required instruments in each situation. We further work on enhancing the system's flexibility by integrating verbal information given by the surgeon. Further user experiments will put a stronger focus on OR staff members, particularly scrub nurses, and their collaboration with the remaining OR team.

References

[1] Carpintero, E., Pérez, C., Morales, R., et al., Development of a robotic scrub nurse for the operating theatre. *3rd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*, 2010

[2] C. Schäfer. Risikofaktor Pflegekräftemangel. *Dialyse aktuell*, 16 (09): 483-483, 2012, in German

[3] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (pp. 91-99).

[4] R. Stauder, N. Navab et al. Learning-based Surgical Workflow Detection from Intra-Operative Signals. *arXiv preprint arXiv:1706.00587*, 2017

[5] Young, T., Hazarika, D., Poria, S., & Cambria, E. (2017). Recent trends in deep learning based natural language processing. *arXiv preprint arXiv:1708.02709*.

[6] C. Kamphuis et al. Augmented reality in medical education? *Perspectives on medical education* 3.4: 300-311, 2014

[7] N. Navab, T. Blum et al. First deployments of augmented reality in operating rooms. *Computer*, 45 (7): 48-55, 2012

[8] B. Laugwitz, M. Schrepp et al. Construction and evaluation of a user experience questionnaire. *Symposium of the Austrian HCI and Usability Engineering Group*, 2008

[9] M. Schrepp et al. User Experience Questionnaire Benchmark Praxiserfahrungen zum Einsatz im Business-Umfeld. *Usability Professionals*, 13: 348-353, 2013, in German

Acknowledgements

The project was supported by Deutsche Forschungsgemeinschaft (DFG, grant FOR 1321)