Christian Kunz, Paulina Maurer, Fabian Kees, Pit Henrich, Christian Marzi, Michal Hlaváč, Max Schneider and Franziska Mathis-Ullrich*

Infrared marker tracking with the HoloLens for neurosurgical interventions

https://doi.org/10.1515/cdbme-2020-0027

Abstract: Patient tracking is an essential part in a surgical augmented reality system for correct hologram to patient registration. Augmented reality can support a surgeon with visual assistance to navigate more precisely during neurosurgical interventions. In this work, a system for patient tracking based on infrared markers is proposed. These markers are widely used in medical applications and meet the special medical requirements such as sterilizability. A tracking accuracy of 0.76 mm is achieved when using the near field reflectivity and depth sensor of the HoloLens. On the HoloLens a performance of 55–60 fps is reached, which grants a sufficiently stable placement of the holograms in the operating room.

Keywords: augmented reality; computer assisted surgery; IR-marker; navigation; tracking.

Introduction

Neurosurgical interventions may be greatly supported through visual guidance by Augmented Reality (AR), potentially leading to a better outcome for the patient. AR allows to superimpose hidden structures over the patient's visible features and visualize them to the surgeon. Only if the patient is tracked with sufficient accuracy, the holograms can maintain the correct position over time. Motion parallax of the AR glasses wearer must be compensated and the patient's movements must be tracked so that the position of the holograms can be updated in relation to the real world. Marker based tracking is a potent and widely-used technology for this purpose. Thus, a great variety of two dimensional markers exist, such as AprilTags [1], Aruco [2], and the commercially distributed product Vuforia [3]. Frantz et al. use Vuforia markers to realize tracking of a patient with the HoloLens [4] in neurosurgical interventions. Other approaches realize patient tracking through surface or contour matching. Wang et al. propose a system for dental surgery using contour matching to track the movements of the patient [5]. Due to the rigid connection of the tooth to the skull, dental procedures are very well suited for tracking via surface matching. For neurosurgical interventions, however, feature based tracking is often not an option, because the patient is fully covered. Infrared (IR) markers are particularly suitable in the medical context and are well established. They provide many advantages, such as sterilizability. Baum et al. proposed a neuronavigation system with the HoloLens, in which the tracking of the patient is realized via an OptiTrack system and IR markers attached to the patient [6]. Kockro et al. pursue a hand-held video probe with a NDI Polaris system for neurosurgical navigation [7] with AR being displayed on a monitor. An alternative approach is proposed by Tabrizi and Mahvash, where the patient is tracked with a Vicon system and information is projected on the patient skin for tumor removal [8]. Most medical technology companies that provide solutions requiring patient tracking develop their own tracking solutions based on IR markers, such as Brainlab or Stryker. In this work a novel approach for IR marker based tracking on the Microsoft HoloLens is presented.

Material and methods

Sensor capabilities of the HoloLens

All sensors of the HoloLens are accessible through the system's research mode: four environmental cameras (*ll*, *lf*, *rf*, *rr*), one near field Time of Flight (ToF) sensor (short throw reflectivity (str) and short throw depth (std)), one long field ToF sensor (*long throw reflectivity and depth*) and a RGB front camera. Two different approaches for tracking IR markers with the HoloLens are investigated. On the one hand, tracking can be achieved by mapping the *str* stream to the *std* stream. Another option is to use the two environmental cameras *lf* (left front) and *rf* (right front), which form a stereo camera setup. Here, both

***Corresponding author: Franziska Mathis-Ullrich,** Health Robotics and Automation, Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, 76131, Karlsruhe, Germany, E-mail: franziska.ullrich@kit.edu
Christian Kunz, Paulina Maurer, Fabian Kees, Pit Henrich and Christian Marzi, Health Robotics and Automation, Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, 76131, Karlsruhe, Germany, E-mail: christian.kunz@kit.edu (C. Kunz), uindc@student.kit.edu (P. Maurer)
Michal Hlaváč and Max Schneider, Department of Neurosurgery, University of Ulm, Günzburg, Germany

approaches are evaluated in regard to their achievable tracking accuracy. The IR markers are brightly visualized in the *str* stream, even when several meters away from the HoloLens. In the streams of the environmental front cameras the IR markers are visible but not brighter than their surroundings. This makes it difficult to create a binarized frame, in which only the IR markers are visible, whereas in the *str* stream this is easily possible. To make the IR markers glow in the *lf* and *rf* streams, they are illuminated with an additional IR light source. Therefore, an IR light emitter was constructed and mounted on the HoloLens as depicted in Figure 1. The LEDs used provide a radiometric power of 530 mW and emit IR-radiation at a wavelength of 850 nm. With the additional illumination it is possible to perceive the IR markers in the environmental cameras under low light conditions. This possibility was examined as tracking of objects through the front cameras could theoretically exhibits a better precision due to higher resolution and better image quality. The *lf* and *rf* frames have a resolution of 640 × 480, while the short throw reflectivity and depth frames have a lower resolution of 448 × 450. Additionally, in the *str* and *std* streams, the sensor only captures data in a centrally located circular area. This leads to larger areas of black pixels in the areas around the inner circle of a frame. The sensor streams of the HoloLens are shown in Figure 2. In the following, tracking on the HoloLens is implemented using the *str* and *std* streams. The offline accuracy evaluations are presented for both tracking approaches. On the basis of the tracking accuracy and implementation capability it was decided to implement the tracking on the HoloLens using the *str* and *std* streams.

IR marker tracking process

In every update cycle of the HoloLens, the current *str* and *std* frames are requested. The IR target tracking process is illustrated in Figure 3. Each IR target consists of four IR markers that allow the tracking of the target's pose. First the *str* frame is binarized with a threshold [T=250, 255] using the binarization function 1 performed on each pixel p with value p_{val} of the *str* frame S_r. This value range is chosen as reflective markers cause brightness clipping as a result of their high light reflectance.

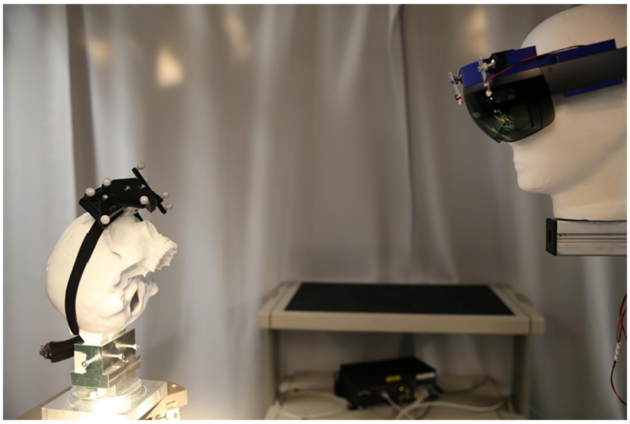

Figure 1: Custom phantom with mounted Infrared (IR) marker system in the experimental setup and aligned HoloLens.

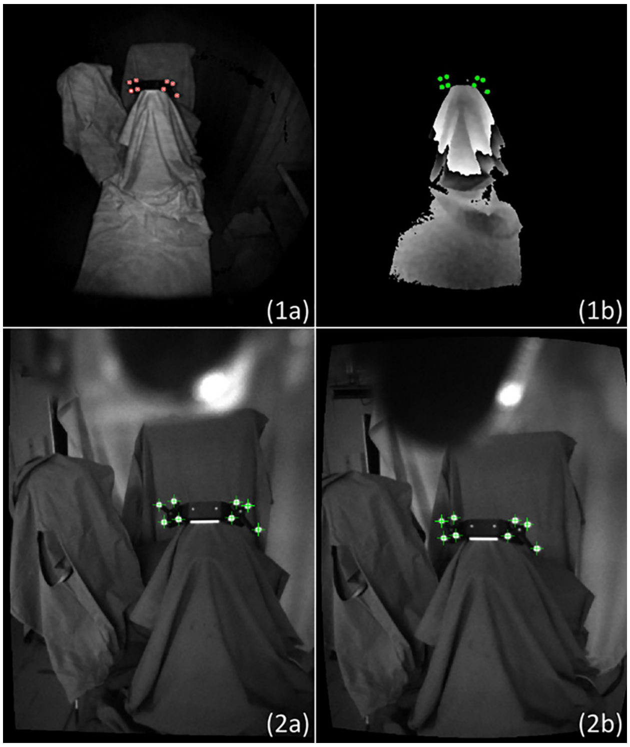

Figure 2: Streams of the HoloLens: (1a) *str* and (1b) *std* with detected blobs, (2a) *lf* and (2b) *rf* with detected blobs.

$$\mathcal{B}: \mathcal{S}_r \rightarrow \{0,1\}, \mathcal{B}(p) = \begin{cases} 1 & p_{val} \in T \\ 0 & else \end{cases} \qquad (1)$$

A blob detection algorithm is then executed to detect the IR markers in each of the binarized frames. In this step any unwanted reflections are filtered. In some instances spherical reflections remain in the image during this step, which are later deleted in the 3D reconstruction phase. The center of each blob is calculated to obtain the 2D position of the IR markers in the *str* frame. The HoloLens offers an unprojection mapping, which allows the retrieval of the 3D positions [X, Y, Z] from every point in the *str* frame. The unprojection

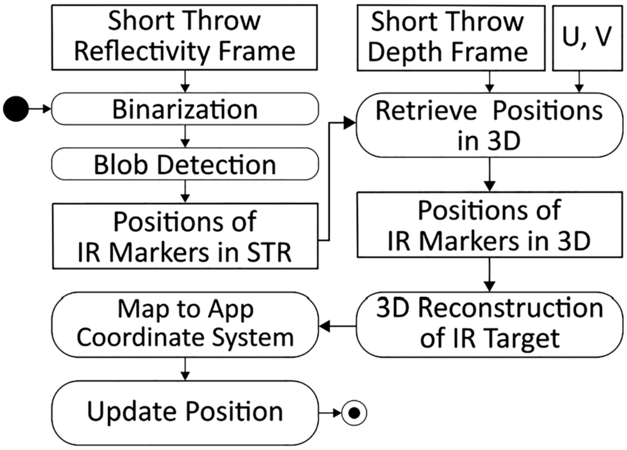
Short Throw Reflectivity Frame
Binarization
Blob Detection
Positions of IR Markers in STR
Map to App Coordinate System
Update Position
Short Throw Depth Frame
U, V
Retrieve Positions in 3D
Positions of IR Markers in 3D
3D Reconstruction of IR Target

Figure 3: Process to determine the pose of the IR targets.

mapping uses two additional matrices U and V, which describe the relationship between points on the depth frame *std* and their 3D positions in the real world. For any pixel Z_{ij}, we can compute the 3D position as follows:

$$[X, Y, Z] = Z_{ij} \cdot [U_{ij}, V_{ij}, 1] \tag{2}$$

This step is described in more detail in [9]. After retrieval of the 3D positions of the retroreflective spheres (see Figure 5, left), the 3D reconstruction is performed to determine the pose of the IR target in space. Since the spatial relationships of the target are known, it is possible to draw lines between opposite IR markers. The intersection point of the lines defines the source of the target. The HoloLens uses a unique coordinate system in every app: *the app specific coordinate space*. Therefore, the retrieved points are mapped to this specific coordinate frame. In the last step the position of the IR targets are updated.

Experiments

To measure the accuracy of the two proposed approaches to track the IR markers, several experiments are conducted and recorded. The IR marker targets are mounted on a custom phantom head placed on a robotic Stewart platform (Hexapod, PI, Germany). The Hexapod's movement accuracy is reported as 1 μm, which enables a very precise ground truth of motion. The platform is moved several times in all three translational axes, x, y, and z. Rotational experiments are not performed in this study. The HoloLens is aligned with the marker system at a distance of 60 cm. Recordings on the HoloLens are collected with the *HoloLensForCV* project [10]. During experimentation, 10 movements for each axis and 10 combined movements across all axes were performed. Both variants of the IR tracking algorithms are then evaluated offline by analysis of the recordings and determination of the position change of the IR targets in space. The real movement is compared to the results of the algorithms. Comparison is based on the vector length in space. The implemented method on the HoloLens is evaluated by measuring the runtime of each algorithm step and the framerate (fps) that was achieved on the HoloLens.

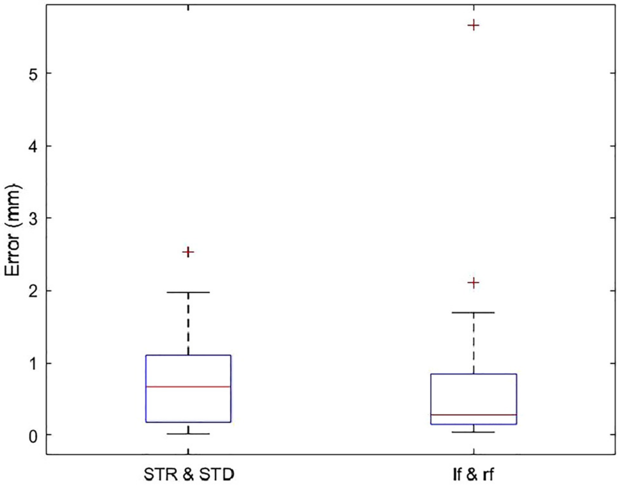

Figure 4: Tracking error for IR target tracking using STR and STD (left) and IR target tracking with *lf* and *rf* (right).

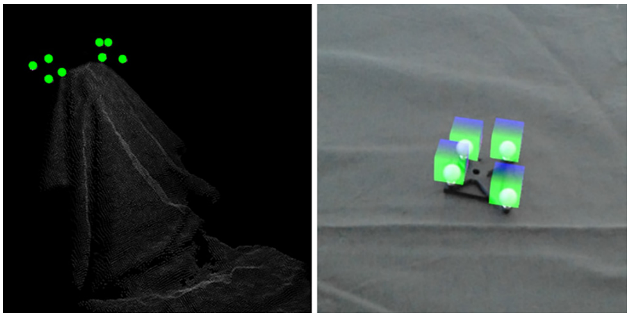

Figure 5: IR markers in reconstructed pointcloud after unprojection (left). IR target visualized in HoloLens (right).

Results

The results depicted in Figure 4 show similar tracking accuracy for both of the two investigated approaches. The average tracking accuracy is 0.76 mm when using *str* and *std* and 0.69 mm for *lf* and *rf*. Results of the marker tracking on a single *lf* and *rf* frame is shown in Figure 2 (1a and b). The centers of the IR markers are detected with sufficient accuracy. Same applies to the *str* frame as shown in Figure 2 (2a and b). Although tracking through *lf* and *rf* yields sligthly better results, it was decided to use the approach using *str* and *std*. The variant utilizing the environmental cameras is dependent on additional IR light emitter sources that need to be mounted on the HoloLens, resulting in additional weight. This is not necessary when using the *str* and *std* resulting in the preferable option. Analysing the results of the performance measurements as depicted in Table 1, 22 frames can be processed with the proposed IR marker tracking algorithm. Performance reserves remain on the HoloLens, since the CPU is only used at 60–70% capacity. A scene from the HoloLens with visualized IR marker target is shown in Figure 5.

Table 1: Results of performance tests divided in all individual process steps.

Process step	Execution time in s
Binarization	0.000583
Blob detection	0.043253
Retrieve positions in 3D	0.000007
3D reconstruction of IR target	0.00103
Map to app coordinate system	0.000001
Update position	0.000001
All	0.044875 (22fps)
Framerate on HoloLens	55–60 fps
CPU utilization	60–70%

Discussion

The results show that tracking of IR targets with the HoloLens is possible with high accuracy of 0.76 mm when using the *str* and *std* frames and 0.69 mm when using *lf* and *rf*. Manual review of the blob detection results show that the IR markers are detected in the *str*, *lf* and *rf* streams with high accuracy and robustness. In addition, IR emitters are necessary to enlight the IR markers in *lf* and *rf* making the HoloLens heavier and handling more difficult. Thus, our approach was the implementation of visual tracking on the HoloLens in combination with the *str* and *std* streams. The good depth sensor quality was confirmed through the good tracking result of 0.76 mm on average (n=42). A performance of 55 up to 60 fps was achieved when executed on the HoloLens. The proposed algorithm can process at least 22 fps of the *str* stream that is limited to 30 fps. The pose of an IR target can be updated every 34 ms (30 fps) to 45 ms (22 fps). The implemented solution uses asynchronous multi-threading that enables the system to process more than 22 fps of the *str* stream, leading to 22–30 processed frames. The *std* stream ist throttled to 23 fps during fast movements limiting the maximum processed fps for a short period of time. A qualitative evaluation on the HoloLens confirms this fast processing speed. In some cases, drift in the translational position is observed during tracking, which will be optimized during future work. Blob detection will also be evaluated, as it requires most of the total processing time and can possibly be optimized in further work.

Conclusion

In this work we propose a solution to track IR marker targets with the HoloLens. Hereby, two different approaches are implemented offline and compared to each other. Tracking is performed using the *str* in combination with the *std* stream or through the two environmental cameras *lf* and *rf*. The first approach yields a tracking accuracy of 0.76 mm at a distance of 40–60 cm. This approach was implemented on the HoloLens reaching high framerates of 55–60 fps. The results enable our system to use well established and sterilizable medical IR markers for AR in the operating room when tracking the relative motion between a patient and the HoloLens.

Research funding: The author state no funding involved.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Competing of interests: Authors state no conflict of interest.
Informed consent: Informed consent was obtained from all individuals included in this study.

References

1. Wang J, Olson E. Apriltag 2: efficient and robust fiducial detection. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, Daejeon, South Korea; 2016.
2. Garrido-Jurado S, Muñoz-Salinas R, Madrid-Cuevas FJ, Marín-Jiménez MJ. Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recogn 2014;47:2280–92.
3. PTC. Vuforia marker; 2020. Available from: https://www.ptc.com/de/products/augmented-reality.
4. Frantz T, Jansen B, Duerinck J, Vandemeulebroucke J. Augmenting microsoft's hololens with vuforia tracking for neuronavigation. Healthc Technol Lett 2018;5:221–5.
5. Wang J, Suenaga H, Hoshi K, Yang L, Kobayashi E, Sakuma I et al. Augmented reality navigation with automatic marker-free image registration using 3-d image overlay for dental surgery. IEEE Trans Biomed Eng 2014;61:1295–304.
6. Baum ZM, Lasso A, Ryan S, Ungi T, Rae E, Zevin B et al. Augmented reality training platform for neurosurgical burr hole localization. J Med Robot Res 2020:194–2001. https://doi.org/10.1142/S2424905X19420017.
7. Kockro RA, Tsai YT, Ng I, Hwang P, Zhu C, Agusanto K et al. Dex-ray: augmented reality neurosurgical navigation with a handheld video probe. Neurosurgery 2009;65:795–808.
8. Tabrizi LB, Mahvash M. Augmented reality–guided neurosurgery: accuracy and intraoperative application of an image projection technique. J Neurosurgery 2015;123:206–11.
9. Labini MS, Gsaxner C, Pepe A, Wallner J, Egger J, Bevilacqua V. Depth-awareness in a system for mixed-reality aided surgical procedures. In: International conference on intelligent computing. Springer, Cham; 2019.
10. Microsoft. Hololensforcv c#/c++ library; 2019. Available from: https://github.com/microsoft/HoloLensForCV.