A Practical AR-based Surgical Navigation System Using Optical See-through Head Mounted Display

Mai Trinh
Department of Computer Science
University of Houston
Houston, USA
mtrinh2@uh.edu

Nikhil V. Navkar
Department of Surgery
Hamad Medical Corporation
Doha, Qatar
nnavkar@hamad.qa

Zhigang Deng
Department of Computer Science
University of Houston
Houston, USA
zdeng4@central.uh.edu

***Abstract*—The work presents a practical Augmented Reality (AR) based surgical navigation system using optical see-through head-mounted display as a standalone solution, without the need of additional tracking hardware. Specifically, we propose a fiducial marker-based instrument tracking, which entirely relies on the built-in hardware of the Microsoft HoloLens 2. The tracking algorithm computes the pose of the tracked object from the real-time image obtained from the on-board front-facing RGB camera. The estimated transformation is then transmitted back to the HoloLens for visualization. Our experimental evaluation shows that the system can achieve 0.81 mm / 1.52 degree in tracking accuracy and sub-millimeter alignment accuracy.**

***Keywords*-Augmented Reality, Surgical Navigation System, Tool Tracking, Surgical Guidance, Head-mounted Displays**

I. Introduction

Surgical Navigation Systems (SNS) are used to accurately, real-time track the positions of surgical instruments with respect to the patient's anatomy. The tracking provides real-time guidance to navigate the surgical instrument during an intervention. Commercial SNS, commonly used in the operating room, utilizes infrared for tracking retro-reflective markers attached to the surgical instrument. The virtual information pertaining to instrument motion with respect to the patient's anatomical landmarks (or preoperative medical images) is displayed on a visualization screen. This causes the surgeon's focus to shift, which may lead to disruption in hand-eye coordination [1]. Though optical camera tracking systems provide a high accuracy and a large tracking area, they suffer from the line-of-sight issue. As the operating room is jam-packed with surgical staff, the markers might get obstructed, and in this case no information is rendered to the surgeon. This limits the mobility of surgical staff in the operating room.

Augmented Reality (AR) can provide a unique solution to the aforementioned limitations as it seamlessly combines physical and virtual environments [2]–[4]. In particular, AR-based Optical See-Through Head-Mounted Display (OST-HMD) devices offer numerous advantages: (i) they allow information in form of virtual objects to be overlaid onto the operator's physical environment; (ii) they alleviate the issue of shifting of the operator's focus and thus reducing the cognitive load [5]. As a result, an AR-based SNS using OST-HMD can significantly enhance surgical navigation. However, existing AR-based SNS still rely on the use of external tracking systems (to track the surgical instruments and the patient) and utilize AR headset devices as a visualization tool only [6], [7]. Since an additional hardware device is required, besides its cost, this would occupy additional space in the operating room and make it more difficult to integrate the additional device into the surgical workflow. Thus, the development of a low-cost, practical, easy-to-integrate, AR-based SNS using OST-HMD (without requiring an additional tracking hardware system) would be beneficial.

In this work we present an AR-based SNS which uses Microsoft HoloLens 2 (a OST-HMD device by Microsoft Corporation, Redmond, USA) as a standalone solution. We describe a fiducial marker-based tracking approach for the tracking of surgical instruments using exclusively the HoloLen's built-in cameras, thus removing the need for external tracking system. A 3D-printed dodecahedron (with 12 attached fiducial markers) attached to an instrument is used to track the poses (position and orientation), and virtual objects are overlaid onto the physical reality in the AR view seen through the HoloLens. As a proof of concept for SNS, we enable landmark registration using a trackable tool to geometrically align the virtual and the physical representations of a spine model.

II. Related Work

Vision-based tracking technologies are one of most prevalent approaches for AR-based medical applications. We briefly review related works that employ tracking for OST-HMDs using different vision-based tracking methods.

A. Marker-less Tracking Approaches

Gu et al. [4] studied the feasibility of a marker-less image based AR for intra-operative surgical guidance during the total shoulder arthroplasty procedure using Microsoft HoloLens 1. Their method utilizes the built-in time-of-flight depth camera for accurate simultaneous indoor localization

2471-7819/22/$31.00 ©2022 IEEE
DOI 10.1109/BIBE55377.2022.00043

and mapping (SLAM) of environment and align a 3D model with the patient's anatomy. Their experimental results show the system does not meet accuracy requirements for clinical applications due to centimeter-level SLAM accuracy and hardware limitation for proper depth sensor technology.

B. IR-based Optical Tracking Approaches

Several works used an external optical tracking system in combination with OST-HMD to develop AR-based surgical guidance/navigation systems [6], [7]. While these systems showed clinically-relevant accuracies, external hardware requirement makes such systems expensive, bulky, and limited with regard to the surgeon's movement. To tackle this limitation, Gsaxner et al. [8] proposed a 6-DOF inside-out tracking algorithm utilizing the built-in tracking cameras of the HoloLens 2. Their trackable instrument contains several retro-reflective marker spheres. Their system can achieve a high accuracy and real time performance. The instabilities of the SLAM system and the limitation in depth sensor technology affect the accuracy of the system, which are common issues when using the built-in depth camera.

C. Marker-based Tracking Approaches

Fiducial marker-based tracking is a popular method used for medical applications. ARsisst [3] is an AR application for assistance in a robot-assisted laparoscopic surgery using OST-HMD. A single fiducial marker is attached to the robotic arm and the hand-held instrument. A hybrid tracking scheme is employed to offer an accurate overlay visualization of the tracking tools. Qian et al. [9] proposed ARAMIS, an AR system providing real-time imaging of the anatomy in the laparoscopic surgery. An endoscope-attached fiducial marker is used to estimate the pose of the endoscope tip in order to display images at correct positions and depths. These systems offer an overlay accuracy of around $4.2 \sim 4.6$ mm, which is substandard in clinical scenarios. The use of a single fiducial marker resulted in a limited range of object motion and a narrow working area.

III. Method

The workflow of our proposed AR-SNS is shown in Fig. 1. Our system runs directly on the HoloLens 2, utilizing the built-in, front-facing, RGB camera to real-time obtain images. The camera records frames of 896 x 504 resolution at approximately 20 Hz. The tracking algorithm processes the images and provides an estimated transformation of the tracked object using the proposed fiducial marker tracking method. After the transformation is obtained, we can estimate the pose of the tip in the world coordinate system. In surgical navigation applications, the operator first navigates the tracked object to each of the pre-determined landmarks on the physical model. The user-selected landmarks are then aligned with the corresponding landmarks in virtual environment. The augmented visualization is displayed in

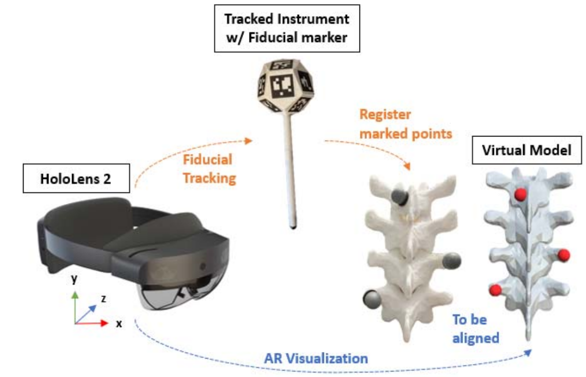


Figure 1. Overview of the proposed AR-based SNS workflow. The user navigates the tracked object and align object's tip to each marked points, while wearing the headset. Real-time images of tracked object are captured by HL2's built-in RGB camera and processed by the tracking algorithm. Final transformation is streamed back and visualized as virtual overlay.

the HoloLens 2 via a Unity-based AR app. The proposed tracking system is implemented in C++ using the OpenCV library [10] and the ArUco library [11], [12].

A. Trackable Instrument Design

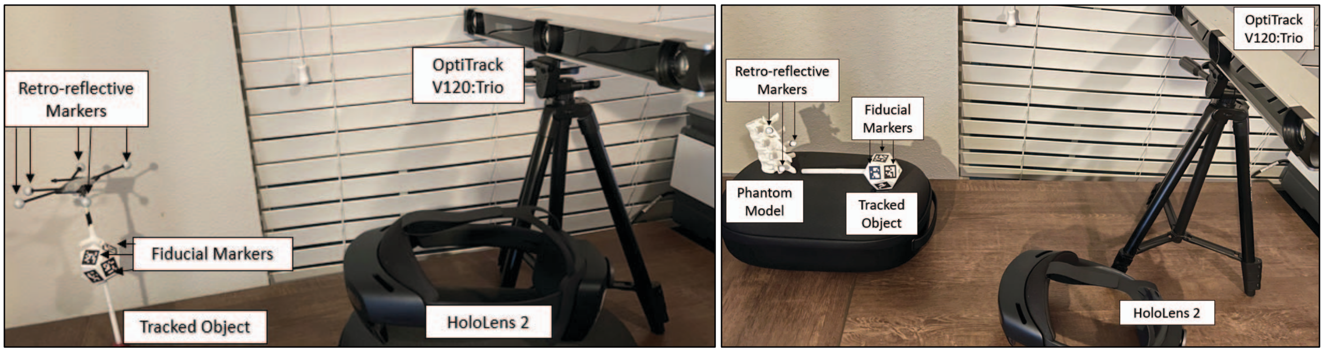
We define a trackable instrument by a set of 12 square binary markers attached to each face of a 3D-printed dodecahedron design, as shown in Fig. 1. The dodecahedron design is used to address the issue of occlusions when tracking planar markers. Each edge of the dodecahedron is 22 mm in length, and the edge of each marker is 20 mm.

The 3D configuration of the trackable instrument is defined by the relative pair-wise transformation between each marker and the reference marker [13]. This is done by capturing multiple images of the trackable object from different viewpoints and apply an algorithm to find the optimal transformation from each marker to the reference marker [14]. Instrument configuration is performed offline and provided to the object tracker at runtime to ensure efficient tracking.

B. Pose Estimation

The built-in camera of the HoloLens 2 provides an input camera stream of the real-world. The tracking algorithm will then determine the position and orientation of the trackable instrument. The system detects the square binary markers attached to the object, and estimates the pose of each marker by first detecting the four corners of the marker and applying the standard Perspective-n-Point (PnP) algorithm [10]. Using the camera's intrinsic parameters and the configuration of 3D markers, we determine the relative transformation of the reference marker with respect to the camera as long as at least one marker is detected. The estimated pose of the object is represented in terms of translations and rotations.

The Levenberg-Marquardt optimization algorithm is used to minimize the reprojection error [15], and the Kalman Filter is applied to the estimated result for smoothing out the



(a) (b)

Figure 2. Experimental setup to evaluate the accuracy of the proposed surgical navigation system during (a) tracking and (b) alignment.

motion. Our tracking algorithm follows the standard Kalman Filter method which consists of two steps, prediction and correction [16]. The algorithm predicts the 3D object pose at each frame, and refines the prediction every time a new measurement of the object is available. The tracking algorithm is explained in detail in [14].

C. Surgical Navigation

The visualization of the surgical navigation application is built on top of Unity3D and is displayed onto the operator's field of view when wearing the HoloLens 2.

1) Registration: Once the estimated pose of the tracked object is obtained, it is passed to a Unity-based AR app. We then estimate the pose of the tip in the world coordinate system by adding 0.14 m along the direction of the instrument to account for the length of the instrument. Our proposed SNS relies on digitizing landmarks using the trackable tool to geometrically align the virtual model with the real object. The operator points the tracked object to each of the pre-determined landmarks, and confirms the selection using the virtual button. The selected landmarks are then aligned with the corresponding landmarks in the physical environment using least-squares fitting [17]. The final transformation of the virtual model is transmitted to the HoloLens 2 for visualization.

2) Hologram alignment: Since OST-HMDs render 3D objects in their coordinate system in front of the user's view, thus to align the virtual model with the real object, we must obtain the transformation between the virtual scene and the physical world. The transformation T_W^P of the phantom model with respect to the world coordinate system is computed as $T_W^P = T_l \cdot \left(T_V^P\right)^T$, where T_V^P is the transformation of the phantom model in virtual coordinates and T_l is matrix representing the left-handed coordinate system.

IV. Results

We describe the results of our experiments conducted to evaluate the system. Both the tracking accuracy and alignment accuracy validation experiments were done by directly comparing the recorded values with the ground-truth data obtained by an OptiTrack V120:Trio motion capture system (NaturalPoint Inc., USA).

A. Experimental Setup

The experimental setup for evaluation of the tracking accuracy is shown in Fig. 2a. OptiTrack system captures the ground truth pose of the tracked object via the five retro-reflective markers. The tracked object is moved freely by the operator (wearing HoloLens 2) within 0.5 m in front of the camera. Five trials were used to compute the average translation and rotation errors. The experiment setup to assess the alignment accuracy for the SNS is illustrated in Fig. 2b. Three retro-reflective markers are attached to the 3D-printed spine model at pre-determined landmark positions. During the task, the operator aligns the tip of the tracked object with each of the markers. OptiTrack system is used to measure the alignment errors. The alignment errors are computed as the distance between the retro-reflective markers and the tracked object tip.

B. Accuracy of the system

The proposed approach achieves a translation accuracy of 0.81 ± 0.30 mm and a rotation accuracy of 1.52 ± 0.37 degrees. The average alignment errors along x, y, and z direction is 1.07 ± 0.5 mm, 0.39 ± 0.8 mm, and 0.43 ± 0.8 mm, respectively. The hologram alignment achieved by the proposed AR surgical navigation pipeline is presented in Fig. 3. The accuracy is affected by the distance from the camera to the tracked object, thus the optimal results were recorded within the distance ranging from 0.2 m to 0.35 m.

V. Conclusion

In this paper, we present a standalone AR-based SNS using HoloLens2. It relies on digitizing landmark registration to align the virtual model with the models in the physical world. The proposed method (when tested in a controlled environment) can approach the performance of existing SNS (< 1 mm in translation, < 1 degree in rotation) without the need of external tracking hardware. Further user studies would be required to assess performance of the system for navigation [18] and tele-mentoring applications [19], [20].

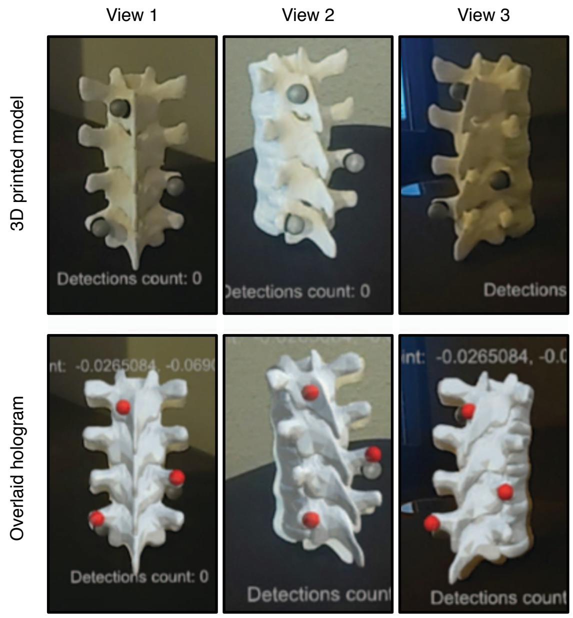


Figure 3. Examples of hologram alignment. **Top:** A 3D-printed spine model with attached optical markers at pre-determined landmark positions for the validation experiment. **Bottom:** The visualization of the hologram alignment from different angles. The virtual spine model with red dots at corresponding positions is overlaid on top of the real 3D-printed model.

ACKNOWLEDGMENT

This work is supported by National Priority Research Program (NPRP) award (NPRP12S-0119-190006) from the Qatar National Research Fund (a member of The Qatar Foundation). All opinions, findings, conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the sponsors.

REFERENCES

[1] R. Wong, J. Jivraj, and V. Yang, "Current limitations and opportunities for surgical navigation," *University of Toronto medical journal*, vol. 92, pp. 7–9, 12 2014.

[2] R. B. Trelease, "From chalkboard, slides, and paper to e-learning: How computing technologies have transformed anatomical sciences education," *Anatomical Sciences Education*, vol. 9, no. 6, pp. 583–602, 2016.

[3] L. Qian, A. Deguet, and P. Kazanzides, "Arssist: augmented reality on a head-mounted display for the first assistant in robotic surgery," *Healthcare Technology Letters*, vol. 5, pp. 194 – 200, 2018.

[4] W. Gu, J. Knopf, N. Navab, and M. Unberath, "Feasibility of image-based augmented reality guidance of total shoulder arthroplasty using microsoft hololens 1," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 9, pp. 1–10, 10 2020.

[5] J. D. Velazco-Garcia, N. V. Navkar, S. Balakrishnan, G. Younes, J. Abi-Nahed, K. Al-Rumaihi, A. Darweesh, M. S. M. Elakkad, A. Al-Ansari *et al.*, "Evaluation of how users interface with holographic augmented reality surgical scenes: Interactive planning MR-Guided prostate biopsies," *Int J Med Robot*, vol. 17, no. 5, p. e2290, Oct 2021.

[6] S. Eom, S. Kim, S. Rahimpour, and M. Gorlatova, "Ar-assisted surgical guidance system for ventriculostomy," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*, 2022, pp. 402–405.

[7] Q. Sun, Y. Mai, R. Yang, T. Ji, X. Jiang, and X. Chen, "Fast and accurate online calibration of optical see-through head-mounted display for ar-based surgical navigation using microsoft hololens," *International journal of computer assisted radiology and surgery*, vol. 15, no. 11, p. 1907—1919, November 2020.

[8] C. Gsaxner, J. Li, A. Pepe, D. Schmalstieg, and J. Egger, "Inside-out instrument tracking for surgical navigation in augmented reality," in *ACM VRST'21*, 2021, pp. 1–11.

[9] L. Qian, X. Zhang, A. Deguet, and P. Kazanzides, "Aramis: Augmented reality assistance for minimally invasive surgery using a head-mounted display," in *MICCAI 2019*, 2019, p. 74–82.

[10] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[11] F. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and Vision Computing*, vol. 76, pp. 38–47, 06 2018.

[12] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and R. Medina-Carnicer, "Generation of fiducial marker dictionaries using mixed integer linear programming," *Pattern Recognition*, vol. 51, pp. 481–491, 10 2015.

[13] H. Sarmadi, R. Muñoz-Salinas, M. A. Berbís, and R. Medina-Carnicer, "Simultaneous multi-view camera pose estimation and object tracking with squared planar markers," *IEEE Access*, vol. 7, pp. 22 927–22 940, 2019.

[14] M. Trinh, J. Padhan, N. V. Navkar, and Z. Deng, "Preliminary design and evaluation of an interfacing mechanism for maneuvering virtual minimally invasive surgical instruments," in *Proc. of IEEE International Symposium on Medical Robotics (ISMR) 2022*, 2022, pp. 1–7.

[15] A. Ranganathan, "The levenberg-marquardt algorithm," *Tutoral on LM algorithm*, vol. 11, no. 1, pp. 101–110, 2004.

[16] G. Welch and G. Bishop, "An introduction to the kalman filter," *Proc. Siggraph Course*, vol. 8, 01 2006.

[17] K. Arun, T. Huang, and S. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 698 – 700, 10 1987.

[18] J. D. Velazco-Garcia, N. V. Navkar, S. Balakrishnan, J. Abi-Nahed, K. Al-Rumaihi, A. Darweesh *et al.*, "End-user evaluation of software-generated intervention planning environment for transrectal magnetic resonance-guided prostate biopsies," *Int J Med Robot*, vol. 17, no. 1, pp. 1–12, Feb 2021.

[19] D. Shabir, N. Abdurahiman, J. Padhan, M. Trinh, S. Balakrishnan, M. Kurer, O. Ali, A. Al-Ansari, E. Yaacoub *et al.*, "Towards development of a tele-mentoring framework for minimally invasive surgeries," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 17, no. 5, p. e2305, 2021.

[20] D. Shabir, N. Abdurahiman, J. Padhan, M. Anbatawi, M. Trinh, S. Balakrishnan, A. Al-Ansari, E. Yaacoub, Z. Deng *et al.*, "Preliminary design and evaluation of a remote tele-mentoring system for minimally invasive surgery," *Surgical endoscopy*, vol. 36, no. 5, pp. 3663–3674, 2022.