Image-based Live Tracking and Registration for AR-Guided Liver Surgery Using Hololens2: A Phantom Study

1st Serouj Khajarian
University of Applied Sciences Landshut
University of Freiburg
Landshut, Germany
serouj.khajarian@haw-landshut.de

2nd Stefanie Remmele
University of Applied Sciences Landshut
Landshut, Germany
stefanie.remmele@haw-landshut.de

3rd Oliver Amft
University of Freiburg
Hahn-Schickard
Freiburg, Germany
amft@cs.uni-freiburg.de

Abstract—We investigate AR-based tracking and registration of the liver surface for potential surgical applications. Our approach consists of streaming RGBD data from a Hololens2 device, RGBD segmentation using a deep learning model and registering the acquired partial liver surface point cloud with the corresponding virtual liver model. We aim to derive basic requirements for AR-guided liver surgery, thus consider several test cases of partially occluded liver as it would appear in surgical scenarios. To evaluate our approach, we use a 3D-printed phantom with basic texture and rigid structure. Our results show that the visible liver section has a substantial impact of feature extraction and matching, thus the registration process. Test cases, where specific anatomical features are visible, e.g. the right liver lobe, yielded superior outcomes compared to other cases, e.g. only the left liver lobe visible. Moreover, our results showed that large scale Hololens movements during the tracking process affected the registration performance. Our implementation achieved 2-3 frames per second for tracking and registration. We discuss the potential and limitations of utilizing Hololens2 for real-time tracking and registration of the liver surface. To our knowledge this is the first experimental approach for real-time markerless tracking and registration for AR-guided surgery guidance using the Hololens2 sensors only.

Index Terms—Augmented Reality, Tracking, liver surgery, Registration

I. Introduction

Numerous studies have investigated how Augmented Reality (AR) could be implemented into surgical procedures, e.g. see [1], [2]. Applications in orthopedic surgeries, in particular, have demonstrated the benefits of AR, as it provides additional visualization capabilities that enhance procedural efficiency and reduce complications and surgery time [3], [4]. However, AR guided navigation in open surgeries, e.g., abdominal surgeries is limited, compared to other surgical applications. A particular challenge for open surgery AR are non-rigid structures and limited visibility of organs during the surgery [1], [5].In liver resection surgery, for instance, the successful implementation of an AR-guided system necessitates continuous tracking of the visible liver surface and the registration of the patient's virtual liver model on the detected surface. The variability of the detected surface poses a challenge in accurate registration of the virtual liver-model

onto its real counterpart, as it limits the use of optical or electromagnetic markers, which are usually used for marker-based tracking and registration methods [5]. In contrast, optical depth sensors (e.g. RGBD-cameras) as used in AR-devices like tablets, smartphones and Head-Mounted-Devices (HMD) combined with advanced algorithms for RGBD-tracking and segmentation have shown promise in markerless tracking of moving objects, particularly in the domains of autonomous driving and AR-supported surgeries [6], [7]. HMDs, in particular, offer intriguing possibilities for AR-guided surgery by enabling in-situ visualization, where relevant information is superimposed directly onto the surgical scene, facilitating an uninterrupted line of sight between the surgeon and the target organ [2], [8]. Nevertheless, the potential of HMDs for live, fully automatic tracking and registration of partially visible target organs without additional tracking sensors has not been thoroughly studied.

In this paper, we asses the feasibility of using an RGBD-tracking algorithm coupled with rigid registration using the integrated sensors of the Hololens2 for tracking the surface of a 3D-printed liver phantom in different environments. We consider the most simple case with easily traceable textures and a rigid organ to explore minimum requirements for the visible part of the liver based solely on sensor quality and liver anatomy. Across the test cases considered, the liver structure was only partially visible, to simulate a realistic surgery situation. Additionally, we determine the computational time requirements for further corrections, such as non-rigid registration and post-processing tracking predictions, which are influenced by the temporal resolution of the AR sensors.

II. Materials and Methods

Our approach to AR-guided liver tracking and registration consists of three steps. **(A) - Data streaming:** Continuous data streaming from the Holelens sensors to a main computer for data processing. **(B) - RGBD-segmentation:** Frame-per-frame segmentation of the visible liver surface using a deep learning model for RGBD-segmentation **(C) - Registration:** Automatic registration of the virtual liver model (created from

prior CT-scan) to the segmented partial liver surface, using a rigid registration algorithm based on feature matching. Our approach is depicted in Fig. 1.
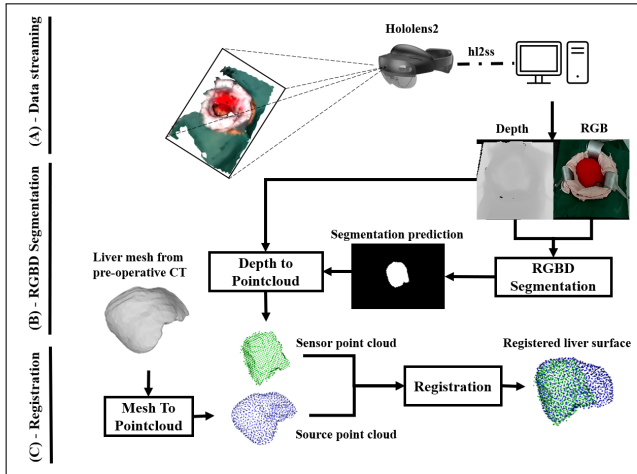

Fig. 1. Overview of our approach to AR-guided liver tracking and registration: (A) Data streaming from the Hololens2 to a main computer. (B) RGBD-segmentation using a deep learning network and generation of the liver surface point cloud (sensor point cloud). (C) Automatic feature matching and registration of the sensor point cloud with the point cloud of the liver mesh (source point cloud).

A. Data streaming

To facilitate the continuous streaming of data from the HMD's RGB camera and time-of-flight based depth sensor to a computer, we employ the Hololens2 sensor streaming application (hl2ss). The application implements a live data transfer via a WLAN connection [9]. The RGB camera captures high-resolution images at 1920x1080 pixels at 30 frames per second (FPS). In contrast, the depth sensor operates at a lower resolution of 320x288 at 1 to 5 FPS. Using extrinsic and intrinsic parameters of both sensors (e.g. sensor position and field of view), we mapped RGB images to corresponding depth images.

B. RGBD segmentation

We utilize a SegNet architecture [10] for tracking the visible part of a red 3D-printed rigid liver phantom surface. The model was trained with previously collected data using the same hardware and phantom. The acquired and adapted RGBD-pair at each frame, serves as the input for the RGBD segmentation network. Upon generating a segmentation prediction, we generated a point cloud representation of the partial liver surface (sensor point cloud).

C. Registration of liver point cloud and model

The partial liver point cloud (sensor point cloud) and its corresponding virtual model were registered to represent the complete liver. we initially create a point cloud comprising 1500 points sampled from the liver model using the Poisson Disk Sampling algorithm (model point cloud). The density of the point cloud was adjusted to resemble that of a fully

detected liver surface acquired using the Hololens sensors, with a voxel size of 5 mm. Subsequently, Fast Point Feature Histogram (FPFH) features [11] are computed for both the sensor and the source point cloud. A global RANSAC (Random Sample Consensus) registration [12] is then performed by feature matching between the two point clouds. Following the global registration, a refinement process is conducted using the Iterative Closest Point (ICP) algorithm [13]. Given the dynamic nature of the generated liver surface point cloud, FPFH features are recomputed after each frame capture. The newly calculated global registration transformation is only applied if the Root Mean Square Error (RMSE) value of correspondences from both point clouds is reduced following the feature computation. In cases where the RMSE value does not decrease, only the registration transformation from the ICP algorithm is employed. All algorithms for the processing and registration of point clouds were implemented using the Open3D Python library [14].

D. Evaluation methods

In order to investigate the minimal visibility requirements for the liver based on the sensor capabilities of the Hololens2 device and the anatomical characteristics of the liver, a series of test cases were created. The test cases involve placing a liver phantom on a table, partially covered by a white cloth, following the scenarios of commonly performed hepatic resection surgeries as outlined in [15]. For the test cases, the liver was only partially visible, to simulate a realistic surgery situation. Additionally, as a reference and benchmark, we performed a full liver surface registration, without occlusions. Furthermore, to simulate a realistic surgical environment, the liver phantom was embedded within a torso phantom featuring a longitudinal incision across the abdomen, allowing for the opening of the abdominal cavity. For our experimental purposes, a red 3D-printed rigid liver phantom is employed. The phantom is based on a segmented CT-scan obtained from the 3Dircad dataset [16]. The color of the phantom was specifically chosen to optimize the tracking performance. A visual depiction of the test cases can be found in Table I.

For each of the created test cases, a 2-minute recording session was conducted, repeated over 5 iterations. Each recording session commenced with the liver being viewed from the visible side (e.g., right side for the test case right hepatectomy), with a rotation of 90° around the table after 35 s to capture a frontal view of the liver. In the final 35 s of the recording, the liver was viewed from the covered side of the respective test case, following another 90° rotation around the table.

To determine the registration performance, we manually annotated all frames with feature matching failures N_{Fail}. We subsequently derived a success rate by $(N-N_{\text{Fail}})/N$ with N being the total number of recorded frames. Furthermore we categorised the calculated success rates according to the four stages of the recordings, consisting of three static viewing points (left view, front view, and right view) and two rotation phases (moving camera). For further analysis, we select the test cases with the best success rates and calculate the minimal

distances between the sensor point cloud and the registered liver meshes. We analyse the median of the calculated distances across the recorded frames and annotate the start and end of both rotation phases in each recording. Additionally, we calculated distances of manually selected correspondent points on randomly selected frame from the reference recordings for a more quantitative representation of the registration accuracy.

III. RESULTS

Table I presents an overview of the success rates, expressed as percentages, for feature matching in the various test cases alongside the average number of acquired frames per recording. The number of frames recorded in a span of 2 minutes per test-case ranged between 250 and 326, indicating that the required computation was achieved at 2-3 frames per second. Notably, test cases with a smaller visible portion of the liver exhibited an approximately 30% higher average number of acquired frames. The highest success rates in feature matching were observed in the test cases *extended right hepatectomy* and the *reference* case, in which the liver was fully visible, followed by the test-case *right hepatectomy*. However, in the *left hepatectomy* and *left lobectomy* test cases, the visible portion of the liver was insufficient to achieve feature matching using FPFH-features of the generated point clouds. For the *torso-phantom* test case, we achieved a success rate of 91% across all recorded frames. Fig. 3 shows examples of successful and failed feature matching.

Fig. 2 depicts median and interquartile range of distances between the generated point clouds (sensor point cloud) and the registered liver meshes in four test cases with the highest feature matching success rates. Median distances ranged between 0.3 mm and 6.5 mm for all the analysed test-cases. Sudden increases in the calculated distances were primarily observed during the head rotation phases of the recordings. The average RMSE value calculated with manually selected corresponding points (3 per frame) from 25 randomly picked frames from the reference recordings was 13 ± 4.8 mm.

IV. DISCUSSION

Our results demonstrate the feasibility of employing the Hololens2 for liver surface tracking and feature matching-based rigid registration during open liver surgery, achieving a frame rate of 2-3 FPS. The frame rate variability was influenced by the number of points of the generated point cloud (sensor point cloud) and the corresponding computational complexity, as shown in Tab. I. Reducing the point density of the sensor and source point clouds resulted in a decrease of feature matching success rate. The achieved frame rate shows the potential for further developments, including post-processing segmentation results or incorporating non-rigid registration techniques similar to [7] to achieve live ($< 1s$) tracking and registration.

The success rates of the feature matching algorithm show a correlation with occlusion level and visible liver anatomy. Specifically, test cases where the larger right lobe of the liver was fully visible exhibited success rates above 95%, while
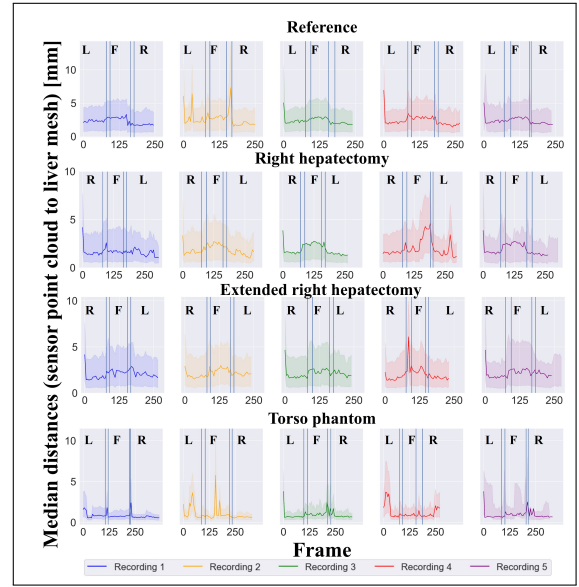

Fig. 2. Distances between sensor point cloud and liver mesh across all test cases: reference, right hepatectomy, extended right hepatectomy, and torso phantom. Solid line: median. Shaded ranges: interquartile range. Letters {R, L, F} identify static phases right view, left view, and front view in the recordings, respectively.
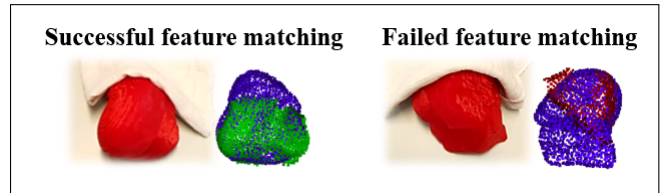

Fig. 3. Example of the outcomes of a successful and failed registration based on matching computed Fast Point Feature Histogram (FPFH) features.

cases with only the left lobe visible resulted in failures. A comparative analysis of the test case *extended left hepatectomy* and *left hepatectomy* highlights an increase in the calculated success rate, which can be attributed to the expanded visible surface area of the liver. Conversely, comparing the test case *extended left hepatectomy* with the test case *right hepatectomy* demonstrates that the enhanced visibility of the right lobe substantially contributes to the improved success rate of feature matching despite the relatively smaller visible surface area (compare Tab. I). The test case involving the torso phantom presents a more realistic scenario, where the liver anatomy is covered from all sides (right lobe was more visible compared to the left lobe). Future studies should analyze additional real-case scenarios using the torso phantom, to further investigate of the applicability and performance of the proposed methodology.

While the calculated distances between the generated point clouds and the corresponding registered liver mesh do not serve as a direct measure of registration accuracy, yet when combined with manually annotated successful feature matching, they provide an approximation of the Hausdorff-distance and the stability of the frame-by-frame registration pro-

TABLE I

<small>THE TABLE SUMMARIZES FRAME COLLECTION, FEATURE MATCHING SUCCESS RATES, AND VISUAL SCENARIOS FOR EACH TEST CASE BASED ON [15].</small>

Test cases		Reference	Right hepatectomy	Left hepatectomy	Extended right hepatectomy	Left lobectomy	Extended left hapatectomy	Torso phantom
Average number of frames per recording		250	292	322	254	354	290	325
Left view (static)		91.51	100	9.38	100	0	38.27	77.72
Right view (static)	Success rate in %	100	95.96	14.387	96.3	0	75.95	99.2
Front view (static)		100	95.04	3.66	100	0	53.5	94.85
Rotating (moving camera)		100	94.32	0	98.57	0	67.97	91.38
Total		97.88	96.33	6.86	98.7175	0	58.92	90.79

(Row header group "Recording phases" spans the four view rows and Total.)

cess (see Fig. 2). Large scale head movements while wearing the Hololens can impact registration stability, as indicated by the increase in calculated distances during the rotation phases. However, normal head movements during static phases do not directly affect registration stability. Although the calculated RMSE value of 13 ± 4.8 mm is in the same range of calculated registration errors of other works, which solely use Hololens for registration [2], the process of selecting correspondences manually requires optimisation. In future work, the registration accuracy should be calculated using clearly identifiable embedded fiducials in the liver phantom.

It is worth noting that the utilization of a unique-color phantom in our experiments offers several advantages, including facilitating label generation and providing realistic depth features. These factors contribute to maximizing the performance of segmentation algorithms. This work specifically aimed to investigate the characteristics of the sensor point cloud, which are influenced by both, the sensor properties and liver anatomy, along with the minimum requirements for real-time tracking and registration. In real-world experiments, the liver shows lower contrast compared to the surrounding tissue, potentially leading to artifacts such as a reduced number of liver points or false positives, which can limit the accuracy of registration. On the other hand, including texture information could potentially provide more features for the registration process. Leveraging texture information to enhance the tracking and registration performance is subject to investigation in future studies.

REFERENCES

[1] Benish Fida, Fabrizio Cutolo, Gregorio di Franco, Mauro FerrariVincenzo Ferrari, "Augmented reality in open surgery," *Updates Surg*, vol. 70, no. 3, pp. 389–400, Sept. 2018.

[2] Longfei Ma, Tianqi Huang, Jie WangHongen Liao, "Visualization, registration and tracking techniques for augmented reality guided surgery: a review," *Phys. Med. Biol.*, vol. 68, no. 4, pp. 04TR02, Feb. 2023, Publisher: IOP Publishing.

[3] Jens T. Verhey, Jack M. Haglin, Erik M. VerheyDavid E. Hartigan, "Virtual, augmented, and mixed reality applications in orthopedic surgery," *Int J Med Robot*, vol. 16, no. 2, Apr. 2020.

[4] Umile Giuseppe Longo, Sergio De Salvatore, Vincenzo Candela, Giuliano Zollo, Giovanni Calabrese, Sara Fioravanti, Lucia Giannone, Anna Marchetti, Maria Grazia De MarinisVincenzo Denaro, "Augmented Reality, Virtual Reality and Artificial Intelligence in Orthopedic Surgery: A Systematic Review," *Applied Sciences*, vol. 11, no. 7, pp. 3253, Apr. 2021.

[5] Tomoyoshi Okamoto, Shinji Onda, Katsuhiko Yanaga, Naoki Suzuki-Asaki Hattori, "Clinical application of navigation surgery using augmented reality in the abdominal field," *Surg Today*, vol. 45, no. 4, pp. 397–406, Apr. 2015.

[6] Ilias Papadeas, Lazaros Tsochatzidis, Angelos AmanatiadisIoannis Pratikakis, "Real-Time Semantic Image Segmentation with Deep Learning for Autonomous Driving: A Survey," *Applied Sciences*, vol. 11, no. 19, pp. 8802, Jan. 2021, Number: 19 Publisher: Multidisciplinary Digital Publishing Institute.

[7] Nicolas Golse, Antoine Petit, Maïté Lewin, Eric VibertStéphane Cotin, "Augmented Reality during Open Liver Surgery Using a Markerless Non-rigid Registration System," *J Gastrointest Surg*, vol. 25, no. 3, pp. 662–671, Mar. 2021.

[8] Mitchell Doughty, Nilesh R. GhugreGraham A. Wright, "Augmenting Performance: A Systematic Review of Optical See-Through Head-Mounted Displays in Surgery," *J. Imaging*, vol. 8, no. 7, pp. 203, July 2022.

[9] Juan C. DibeneEnrique Dunn, "HoloLens 2 Sensor Streaming," Nov. 2022, arXiv:2211.02648 'unpublished'.

[10] Vijay Badrinarayanan, Alex KendallRoberto Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," Oct. 2016, arXiv:1511.00561 [cs].

[11] Radu Bogdan Rusu, Nico BlodowMichael Beetz, "Fast Point Feature Histograms (FPFH) for 3D registration," in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 3212–3217, ISSN: 1050-4729.

[12] Martin A. FischlerRobert C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[13] S. RusinkiewiczM. Levoy, "Efficient variants of the ICP algorithm," *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152, 2001, Conference Name: Third International Conference on 3-D Digital Imaging and Modeling ISBN: 9780769509846 Place: Quebec City, Que., Canada Publisher: IEEE Comput. Soc.

[14] Qian-Yi Zhou, Jaesik ParkVladlen Koltun, "Open3D: A Modern Library for 3D Data Processing," Jan. 2018, arXiv:1801.09847 [cs].

[15] Robert J. AragonNaveenraj L. Solomon, "Techniques of hepatic resection," *J Gastrointest Oncol*, vol. 3, no. 1, pp. 28–40, Mar. 2012.

[16] Luc Soler, Hostettler, Agnus, Charnoz, Fasquel, Moreau, Osswald, BouhadjarMarascaux, "3D image reconstruction for comparison of algorithm database: A patient specific anatomical and medical image database." IRCAD, Strasbourg, France, Tech. Rep (2010)," .