



Augmented reality system for endoscopic pituitary surgery with automatic registration

Aure Enkaoua¹ · João Ramalhinho¹ · Mobarakol Islam¹ · Hani J. Marcus^{1,2} · Matthew J. Clarkson¹

Received: 17 February 2025 / Accepted: 31 March 2025 / Published online: 11 May 2025
© The Author(s) 2025

Abstract

Purpose Endoscopic pituitary surgery is a minimally invasive technique to remove pituitary tumours through the nose. Currently, image guidance may be used in the form of a tracked pointer to help surgeons navigate the region and avoid damage to critical structures. However, the pointer method is mentally demanding as the pointer location is displayed in a different modality and disrupts the surgical workflow due to the setup time and frequent tool removal.

Methods We propose an Augmented Reality (AR) system where information from the pre-operative scan is displayed directly onto the endoscopic video. Our system features an on-board tracking system, allowing for the registration process to be performed automatically.

Results We evaluated the accuracy of our system and compared it to an AR system that uses an infrared (IR) camera to track an endoscope with reflective markers. Our system gave an accuracy of 1.1 (\pm 0.4) mm, compared to 2.4 (\pm 0.9) mm in the IR-tracked endoscope approach.

Conclusion Our Augmented Reality system is a more compact and transportable setup which outperformed the IR-tracked endoscope. The automatic registration method can save time in the operating room as well as increase AR overlay accuracy, improving the translation of these technologies.

Keywords Augmented reality · Transsphenoidal pituitary surgery · Minimally invasive surgery · Endoscopic surgery

Introduction

Endoscopic pituitary surgery is a minimally invasive technique performed to remove pituitary tumours by accessing the base of the skull through the sphenoid sinus with the aid of an endoscope [1]. While it offers advantages over open skull-base surgery [2], it remains high risk due to the close proximity of the pituitary gland to the dense surrounding anatomy.

In order to understand the orientation of the anatomy and avoid damage to critical structures, surgeons currently use a tracked pointer, whose tip location can be synchronously mapped on a pre-operative MRI scan [3]. However, as the

pointer location is shown on a 3D MRI scan but the surgery is seen on a 2D endoscopic video, this method increases cognitive load as the surgeons have to mentally map between different modalities and dimensions. Additionally, the limited nasal cavity space requires the tool to be removed every time a localisation is required, disrupting workflow and delaying the surgery.

Augmented Reality (AR) has been explored as a solution, overlaying pre-operative data on live endoscopic video, with approaches using IR-reflective [4] or visual [5] markers. However, current methods either depend on markers being visible in the pre-operative scan [6, 7], require intra-operative scans to align the patient's anatomy with the navigation system [8], or lack the required precision for pituitary surgery [5] [3]. Fully trackerless systems [9, 10] enable a less complex setup by removing the reliance on external markers. However, they face challenges such as featureless anatomy, limited training data, and hardware constraints, as these techniques would require powerful computational resources in order to work in real time.

✉ Aure Enkaoua
zcemaen@ucl.ac.uk

¹ UCL Hawkes Institute, Department of Medical Physics and Biomedical Engineering, University College London, London, UK

² National Hospital for Neurology and Neurosurgery, London, UK

Methods

We present the first system in endonasal surgery to integrate an on-board tracking system. We compare our system to an IR-tracked endoscope approach, evaluating the overlay accuracy using predefined markers placed in a phantom model.

The phantom used for evaluating both systems was the UpSurgeOn BrainBox TNS model,¹ used for training and simulating an endoscopic pituitary surgery. A CT scan of the phantom was obtained with the Medtronic O-arm CT O2 Intraoperative Imaging System, and its surface segmented and smoothed using 3D Slicer.² A 1 cm grid of spherical markers was placed inside the phantom instead of the replaceable cavity, and the centres of these targets were identified on the CT scan (see Fig. 2). Target models were generated using these locations for the AR overlay. The phantom was held stationary relative to the markers with a custom-made structure simulating a Mayfield clamp. Once set up, both systems followed three key steps: calibration, registration, and tracking.

Our system

Our proposed system can be seen in Fig. 1a. Our setup includes the INTEL® D405 RealSense™ stereo camera³ mounted on a standard Storz endoscope⁴ using a custom 3D printed bracket. An ArUco board was attached to the laser-cut holder that held the phantom stationary.

Calibration: The intrinsic parameters of the endoscope were calibrated using a 5x9 ChArUco board with 20 mm squares, following Zhang's method [11] as implemented in OpenCV [12]. The RealSense™ camera was calibrated using its built-in self-calibration method for the D400 series.⁵ Extrinsic calibration between the cameras was determined by recording frames where the board was visible from both the RealSense™ and endoscope cameras and calculating the relative transform using the two estimated poses.

Registration: As seen on the left of Fig. 1c, to obtain the registration, the phantom was reconstructed using the RealSense camera. Using prior knowledge on the pose of the RealSense camera relative to the face, the segmented CT surface was oriented in the same direction as the reconstructed face. A robust point-to-plane ICP algorithm using Tukey loss ($k = 0.1$) was applied [13] with a distance threshold of 1 mm to mitigate the influence of outliers.

Tracking: Once registered, the 3D model overlay was updated based on the pose obtained from the ArUco board at each frame.

IR-tracked endoscope

The IR-tracked endoscope approach can be seen in Fig. 1b. An NDI⁶ camera was used to track IR-reflective markers attached to the endoscope and phantom.

Calibration: The intrinsic and extrinsic calibration processes were performed using the SmartLiver software as described in [14].

Registration: Seven fiducial markers were placed along the base of the phantom as seen on the right of Fig. 1c, and their locations were localised on the CT scan. During the registration process, these same fiducials were sampled using the tracked pointer tip and could therefore be matched to the CT samples using the procrustes point-based registration algorithm [15].

Tracking: Once registered, the 3D model overlay was updated based on the pose obtained from the patient reference and endoscope reference markers.

Experiment and results

We calculated the projected target registration error (TRE) of our system by measuring the average Euclidean distance between the true target value locations and the AR overlay of the targets. This error includes the calibration, registration, and tracking errors. The experimental setup was designed to mimic clinical conditions, with both systems tested under similar constraints of the endoscope position, lighting, and zoom settings.

The accuracy of the systems was 2.4 (± 0.9) mm and 1.1 (± 0.4) mm for the tracked endoscope and our system, respectively. Figure 2 provides a clear visual representation of the accuracy evaluation, where the true target points are shown in grey and the AR overlay is shown by the red spheres.

Discussion and conclusion

We have presented a new Augmented Reality system for performing pituitary surgery using an on-board tracking system. Our system is more affordable and compact than the current IR-tracked approach, making it more practical for widespread adoption by reducing both cost and the space required in the operating room.

Unlike the approach proposed by Onishi et al. [5], where cameras are placed in the operating room and may be

¹ UpSurgeOn: <https://store.upsurgeon.com/>

² Slicer: <https://www.slicer.org/>

³ RealSense: www.intelrealsense.com/depth-camera-d405

⁴ Karl Storz: <https://www.karlstorz.com/>

⁵ RealSense self-calibration: <https://dev.intelrealsense.com/docs/self-calibration-for-depth-cameras>

⁶ NDI: <https://www.ndigital.com/>

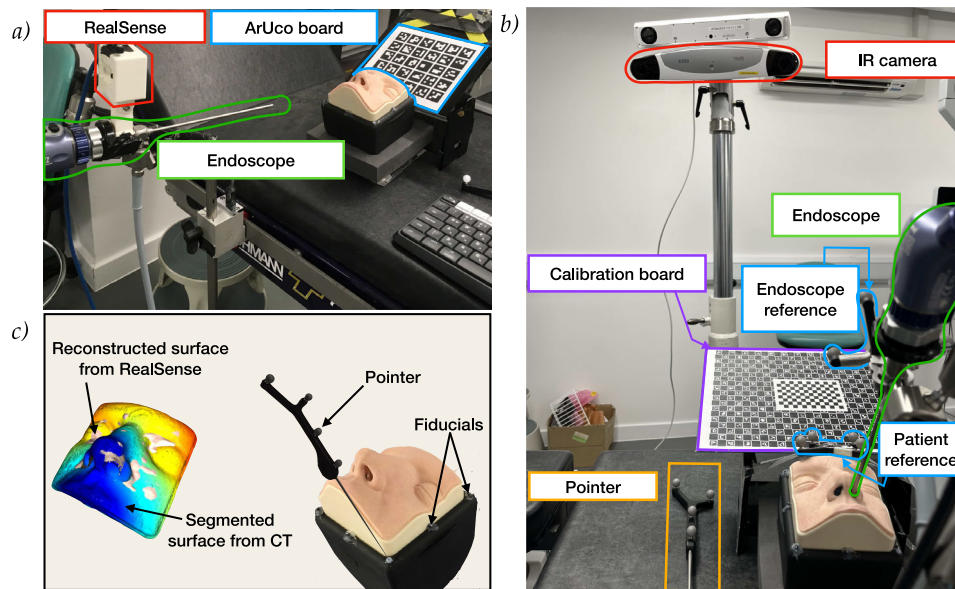


Fig. 1 System overview of our system (a) and IR-tracked endoscope (b). The registration process of our system (left) and the IR-tracked (right) systems is depicted in (c). In our system (a), the ArUco board (blue) acts as the patient reference and is placed on the holder, where the calibration board is also placed. The board is tracked by the external RealSense camera (red), rigidly attached to the endoscope (green). The registration process of our system is shown on the left of (c). The RealSense camera is used to reconstruct the face of the phantom, which

is registered to the CT scan. In the IR system (b), a patient reference (blue) is attached to the phantom holder, and a calibration board (purple) is used to calibrate the endoscope’s (green) intrinsic parameters and its pose relative to the endoscope reference (blue). The IR camera (red) tracks the pointer (orange) and reference markers. Registration as seen on the right of (c) is performed by placing the pointer at seven fiducials with known CT locations for point-based alignment

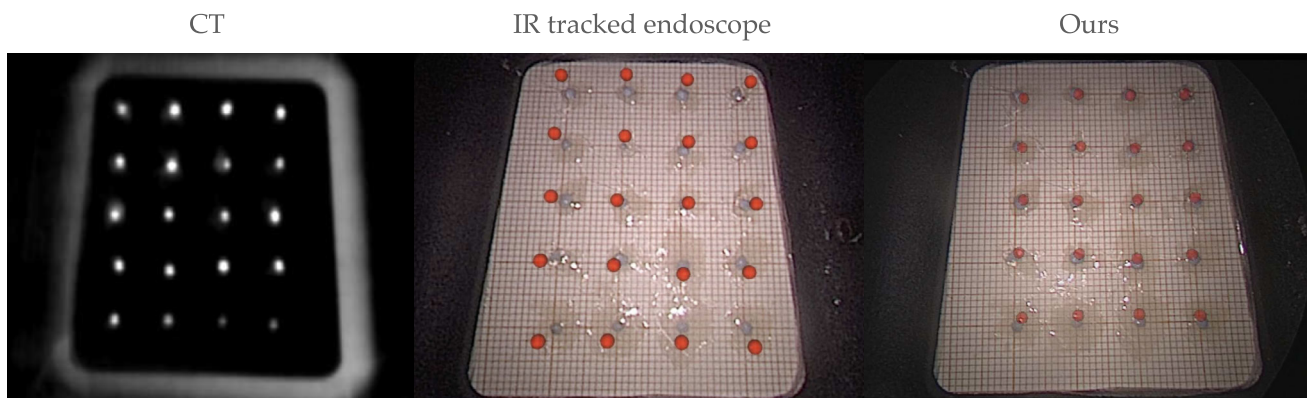


Fig. 2 Accuracy comparison of IR-tracked AR (middle) and our system (right). Grey circles represent true target points, spaced 1 cm apart, while red dots show the AR overlay. The distance between the true

target points (grey) and overlay points (red) represents the accuracy of either system. The CT scan with visible target points used for extracting red 3D locations is shown on the left

obstructed by surgical staff, our system ensures an unobstructed view by mounting the tracking camera directly on the endoscope. Since the endoscope is inserted through the nose, the RealSense camera consistently faces the ArUco board placed on the patient’s head. Because the markers are attached to the patient, they cannot be blocked by surgeons moving around the room. Using an ArUco board instead of

a single tag also allows the system to estimate the pose even when some markers are temporarily occluded.

In previous work, we showed that the IR tracking systems are not accurate enough for the scale of pituitary surgery [3]. Our proposed setup results in a higher accuracy since the tracking camera is placed nearer the reference. Furthermore, while both IR-tracked systems and the method by Onishi et

al. [5] rely on two separate references for tracking, our system requires only a single reference, simplifying the setup.

A key advantage of our system is its automated registration pipeline. Automating this process can reduce setup time and minimise human error, particularly for inexperienced users.

Limitations and future work

Our current evaluation is conducted on a stationary phantom, which allows for a controlled accuracy assessment. Although this study shows promise as a proof of concept, further evaluation is necessary to ensure reliability across different test models and real-world clinical settings.

The registration method used for our system suffers from limitations as it currently relies on a single reconstructed view. Even though this is enough for an initial evaluation with a static environment, it may be sensitive to cases where the reconstruction is poor. To improve robustness against such cases, we plan to explore combining multiple views to the reconstruction, which could fill missing areas in frames where there are occlusions, help detect outliers more effectively, and ultimately enhance the overall accuracy. Additionally, while our current registration method depends on prior knowledge of the expected orientation of the face, a more advanced approach will use real facial landmarks to estimate the initial orientation.

The registration technique used for the IR-tracked endoscope approach, as described in Sect. 2.2, relies on the fiducial markers being visible on the pre-operative scan. However, this method assumes the fiducials will remain fixed between the imaging session and the surgical procedure, which is not practical in a clinical setting. Instead, surface-based registration techniques, where points along the face of the patient are sampled for alignment, are typically used in surgery. Future work will compare the accuracy of our automatic registration and this more realistic registration method.

Funding This work is supported by the EPSRC-funded UCL Centre for Doctoral Training in Intelligent, Integrated Imaging in Healthcare (i4Health) [EP/S021930/1], the Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS) [203145Z/16/Z], and the EPSRC grants [EP/W00805X/1] and [EP/T029404/1]. HJM is supported by the NIHR UCLH/UCL Biomedical Research Centre (BRC).

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethics Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent This article does not contain patient data.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adap-

tation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Marcus HJ et al (2021) Pituitary society expert delphi consensus: operative workflow in endoscopic transsphenoidal pituitary adenoma resection. *Pituitary* 24:839–853
- Komotar RJ, Starke RM, Raper DM, Anand VK, Schwartz TH (2012) Endoscopic skull base surgery: a comprehensive comparison with open transcranial approaches. *Br J Neurosurg* 26:637–648
- Enkaoua A et al (2023) Image-guidance in endoscopic pituitary surgery: an in-silico study of errors involved in tracker-based techniques. *Frontiers in Surgery* 10:1222859
- Li L et al (2016) A novel augmented reality navigation system for endoscopic sinus and skull base surgery: a feasibility study. *PLoS ONE* 11:e0146996
- Onishi K, Fumiyama S, Nonaka M, Koeda M, Noborio H (2021) Study on the Image Overlay Approach to AR Navigation System for Transsphenoidal Surgery. Human-computer interaction. interaction techniques and novel applications: thematic area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part II. Springer International Publishing, Cham, pp 625–643
- Vagdargi P et al (2021) Drill-mounted video guidance for orthopaedic trauma surgery. *J Med Imag* 8:015002–015002
- Kawamata T, Iseki H, Shibasaki T, Hori T (2002) Endoscopic augmented reality navigation system for endonasal transsphenoidal surgery to treat pituitary tumors. *Neurosurgery* 50:1393–1397
- Hajek J, Unberath M, Fotouhi J, Bier B, Lee SC, Osgood G, Maier A, Armand M, Navab N (2018) Closing the Calibration Loop: An Inside-Out-Tracking Paradigm for Augmented Reality in Orthopedic Surgery. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G (eds) Medical image computing and computer assisted intervention – MICCAI 2018: 21st international conference, granada, spain, september 16–20, 2018, Proceedings, Part IV. Springer International Publishing, Cham, pp 299–306
- Espinel Y, Lombion N, Compagnone L, Saroul N, Bartoli A (2024) Preliminary trials of trackerless augmented reality in endoscopic endonasal surgery. *Int J Comput Assist Radiol Surg* 19(7):1385–1389. <https://doi.org/10.1007/s11548-024-03155-6>
- Mirota DJ et al (2011) A system for video-based navigation for endoscopic endonasal skull base surgery. *IEEE Trans Med Imaging* 31:963–976
- Zhang Z (2000) A flexible new technique for camera calibration. *IEEE Trans Pattern Anal Mach Intell* 22:1330–1334
- Itseez. Open source computer vision library. <https://github.com/itseez/opencv> (2015)
- Zhou Q-Y, Park J, Koltun V (2018) Open3D: A modern library for 3D data processing. [arXiv:1801.09847](https://arxiv.org/abs/1801.09847)
- Dowrick T et al (2023) Evaluation of a calibration rig for stereo laparoscopes. *Med Phys* 50:2695–2704

15. Arun KS, Huang TS, Blostein SD (1987) Least-Squares Fitting of Two 3-D Point Sets. *IEEE Trans Pattern Anal Mach Intell* PAMI-9(5):698–700. <https://doi.org/10.1109/TPAMI.1987.4767965>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.