

Maximizing the Accuracy of Ultrasound Imaging with Augmented Reality

Junhao Zhang* and Muyinatu A. Lediju Bell*^{†‡}

*Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD

[†]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD

[‡]Department of Computer Science, Johns Hopkins University, Baltimore, MD

Abstract—Augmented reality (AR) has emerged as a promising solution for freehand ultrasound visualization and guidance in interventional and diagnostic procedures. However, the reliance on cumbersome wired tracking systems often restricts flexibility, and existing AR-based ultrasound navigation systems have tracking errors that limit deployment in high-precision applications. To overcome these limitations, we introduce the first known fully wireless AR ultrasound system that leverages simultaneous localization and mapping algorithms to accurately and efficiently track ultrasound probe locations. We additionally introduce a calibration method that iteratively refines the transformation between the ultrasound probe and the tracker. Our novel system achieves a reprojection error (i.e., mean distance between the same points on two intersecting ultrasound slices in a volume, when projected to the same coordinate system) of 0.983 mm. Results are promising to provide a new, wireless approach to intraoperative assessments, surgical training, and image-guided interventions.

Index Terms—Ultrasound, Augmented reality, Surgical guidance

I. INTRODUCTION

Ultrasound imaging is ubiquitous in modern healthcare, offering real-time, non-invasive visualization for minimally invasive diagnostic and interventional procedures. However, effectiveness can be limited in remote or resource-limited regions and complex surgical settings, as traditional interventional ultrasound systems rely on wired connections [1], require external tracking equipment [2]–[4], and demand mental mapping of 2D ultrasound images to 3D anatomical structures [5].

The integration of augmented reality (AR) with ultrasound imaging has emerged as a promising solution to enhance spatial understanding and procedure guidance [6]–[8]. Recent advances in AR technology, particularly in simultaneous localization and mapping (SLAM), have created new opportunities to overcome traditional limitations in ultrasound image visualization, including real-time spatial alignment of imaging data with patient anatomy, improved depth perception for procedural guidance, and the integration of multimodal medical data in AR environments [9], [10]. These developments coincide with growing interest in wireless medical devices, which offer greater flexibility and improved sterility control in surgical environments.

Despite technological advances in self-localization of AR devices, current solutions depend on complicated external

tracking systems, which can be cumbersome in clinical settings and require additional setup time [11]. Although AR-based marker tracking systems obviate the need for external tracking devices, tracking accuracy is often degraded by image resolution and varied lighting conditions [12]. In addition, the calibration between tracking systems and imaging devices can be challenging when attempting to achieve accurate spatial registration between virtual and real-world elements [13]. Contrasting these challenges with the benefits of ultrasound-guided procedures, a wireless AR-ultrasound system has the potential to enhance efficiency and accuracy.

The work herein presents the first known application of AR-guided wireless ultrasound. We utilize existing SLAM modules from an AR device to track a wireless ultrasound probe without any external markers. This approach enhances flexibility by eliminating wired connections, costly optical tracking cameras, or suboptimal tracking markers that limit clinical viability in traditional AR-based ultrasound-guided interventional procedures. In addition, we propose a novel optimization-based pose refinement method that can be applied to any imaging anatomy, extending beyond the use of a calibration phantom. We demonstrate the feasibility using a human skull.

II. METHODS

A. Framework

A wireless ultrasound probe (Clarius C3HD, Vancouver, Canada) was used to acquire 2D ultrasound slices. An AR device (Magic Leap 2, Plantation, Florida, USA) was utilized to visualize and track the 3D pose of the AR device controller. This controller was securely attached to the ultrasound probe with a custom 3D-printed component (see Fig. 1), enabling probe tracking through the calibration method described in Section II-C.

Data transmission between the server, a computer running a Python script, and the wireless ultrasound probe was implemented using TCP/IP via the PLUS Toolkit [14], which is an open-source library for data communication among image-guided interventional systems. A Unity application (version 2022.3.21f1) was developed to transmit the pose data of the AR device controller to the server, where the pose data were synchronized with the ultrasound image data.

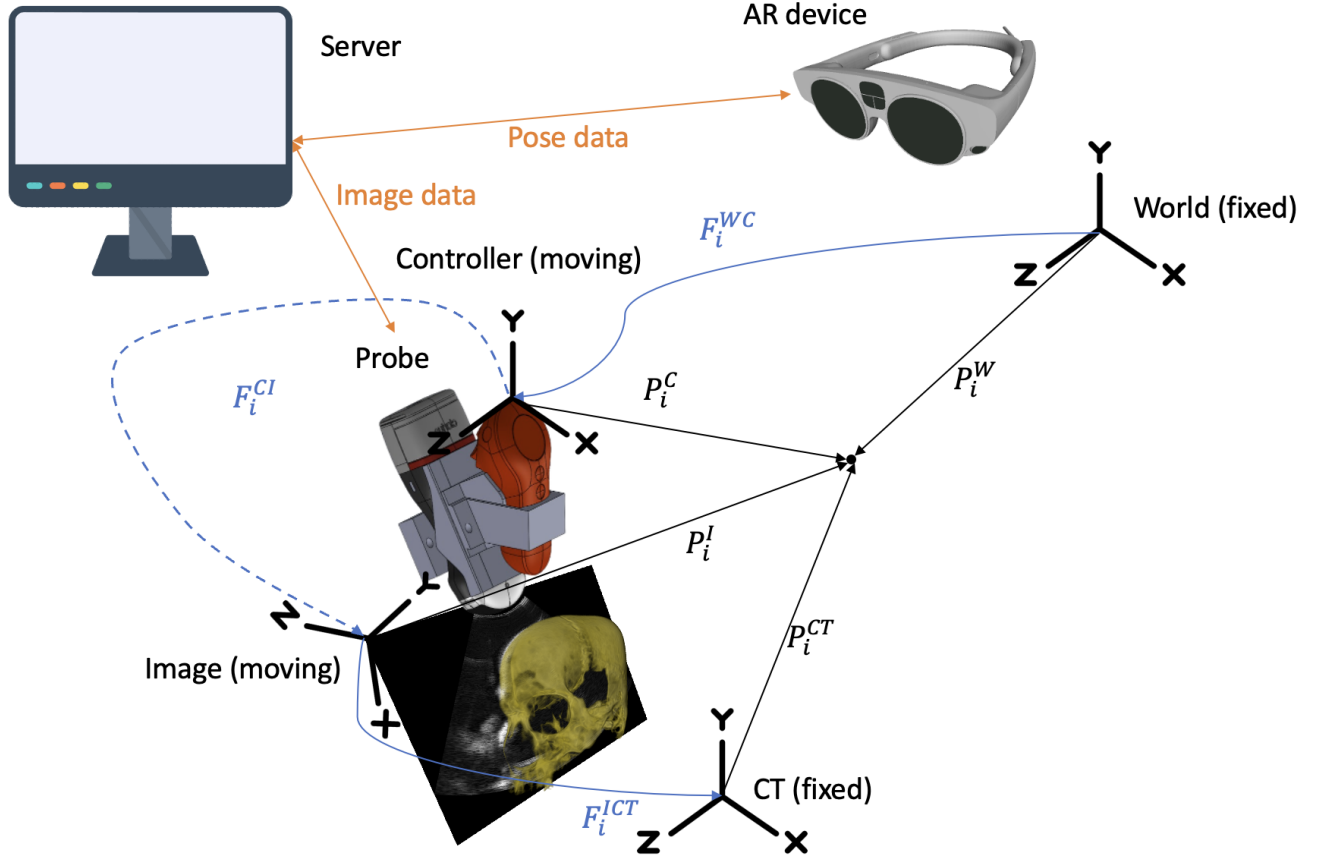


Fig. 1. Major components, coordinate frames and transformations of the proposed system. P_i is the i -th point in the indicated coordinate frame: world (W), controller (C), image (I), or computed tomography of the skull (CT).

B. Coordinate transformations

The coordinate transformations of the system, illustrated in Fig. 1, convert 2D ultrasound image pixels, P_i^P , of the i -th frame into a global world coordinate system, P_i^W , as follows:

$$P_i^W = F_i^{WC} F_i^{CI} F_i^{IP} P_i^P \quad (1)$$

$$= \begin{bmatrix} R_i^{WC} & t_i^{WC} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_i^{CI} & t_i^{CI} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} P_i^P \quad (2)$$

where F_i^{WC} is the transformation matrix from the controller to world coordinate (consisting of a rotation matrix R_i^{WC} and a translation vector t_i^{WC}), F_i^{CI} is the transformation from the image to controller coordinates (consisting of rotation R_i^{CI} and translation t_i^{CI}), F_i^{IP} is the transformation matrix that converts points from the image pixel space to physical coordinates (consisting of scale factors s_x and s_y in the lateral and axial dimensions, which are both 0.362 mm/pixel), and P_i^P is a point in the pixel coordinate of the i -th frame in homogeneous form. In Eq. 1, a 3D pixel (x, y, z) is extended to $(x, y, z, 1)$ as P_i^P or P_i^W to enable linear transformations using matrices

F_i^{WC} , F_i^{CI} , and F_i^{IP} . In addition, P_i^P lies on the set $\{(x, y, 0, 1) \in \mathbb{R}^4 \mid x \in \{0, 1, 2, \dots, 639\}, y \in \{0, 1, 2, \dots, 479\}\}$. The third dimension of P_i^P , representing the elevation axis of the ultrasound probe, is assigned an arbitrary constant value of zero to convert each 2D image coordinate into a 4D homogeneous coordinate, which allows F_i^{WC} and F_i^{CI} to be represented as 4x4 matrices incorporating rotation and translation.

C. Flexible online calibration

Calibration was performed using the concept of a reprojection error, which is the mean Euclidean distance between corresponding point pairs on two intersecting slices, as illustrated in Fig. 2. For each slice in a pair of intersecting 2D slices, the pixel intensity profile along the intersection was computed. The position of the first pixel on the intersection line with an intensity value greater than 200 (out of 255) was identified in each slice and mapped to the world coordinate system. Ideally, these points should coincide if the calibration is correct, resulting in a perfectly tracked probe.

The transformation (in its Lie algebra representation [15]) between the image and controller coordinate systems was optimized with the proposed differentiable reprojection error

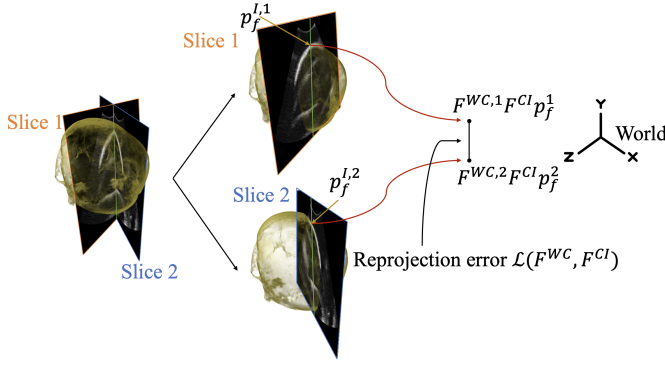


Fig. 2. Reprojection error illustrations.

as a loss function, using the first-order Adam optimizer in PyTorch [16]:

$$F^{CI*} = \arg \min_{F^{CI}} \mathcal{L}(F^{CI}) \quad (3)$$

The learning rate was 0.01 for rotation and 0.001 for translation.

D. Volumetric Reconstruction

A button on the AR controller was programmed to synchronously transmit pose and image data only when pressed. This action also automatically defined the region of interest (ROI) in the physical scene, filtering out any images that did not capture the ROI. Unlike existing volumetric reconstruction toolboxes [17], manual offline ROI selection in image processing software on a computer was not required.

With an elevation slice thickness of 1 mm, 2D ultrasound slices were spatially registered to the global world coordinate using the transformations in Eq. 1. The reconstructed volumetric images had dimensions of 256 x 256 x 256, with voxel sizes varying from 0.3 mm to 0.8 mm, based on the frame rate and scanning field of view. To address gaps in the reconstructed volume, linear interpolation was applied to estimate the intensities of missing voxels. The volumetric images of a skull that was cleaned from tissue attachments (The Phantom Laboratory, Salem, NY, USA), as reported in [18], were visualized using Amira software (Visage Imaging, San Diego, CA, USA).

III. RESULTS & DISCUSSION

Fig. 3 shows an example result pair of the physical setup and the corresponding digital twin, as well as a photograph from a camera placed behind the AR device lens to demonstrate the fusion between a physical and digital twin when viewed from the perspective of a user. Ultrasound images were rendered in the AR application and registered with the real world. The visualization frame rate was 26 Hz, and the average data transmission latency between the server and the AR device was 57 ms.

Fig. 4 shows the reconstructed volumetric ultrasound images of the skull rendered as shaded meshes and compared with the corresponding computed tomography scan as the ground

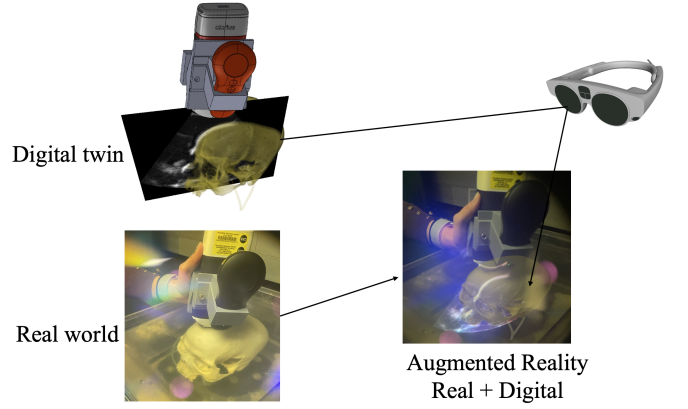


Fig. 3. The real-world physical setup co-registered with virtual elements of the ultrasound slice, X-ray CT, probe, and controller. The AR device processes virtual elements that are overlaid onto the real world when viewed through the AR lens.

truth. Table I reports the reprojection errors obtained with the skull, compared with state-of-the-art AR ultrasound systems employing different tracking methods with internal sensors. Although the specific contexts of reprojection errors vary across the studies reported in Table I [9], [10], [19], each quantifies the average distance between corresponding points within a defined coordinate system. Our proposed differentiable optimization method has the lowest reprojection error.

The mean reprojection error produced by our method, which integrates optimization-based online calibration with an SLAM module, outperforms state-of-the-art approaches that similarly rely on the built-in sensors within an AR device (i.e., the integrated RGB camera [10], [19] or depth camera [9]), as summarized in Table I. In addition, the wireless feature of the ultrasound probe is anticipated to enhance flexibility and simplify the ability of users to maintain direct sight of the

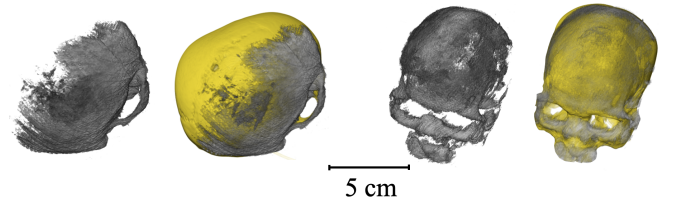


Fig. 4. Co-registered 3D ultrasound (gray) and CT (yellow) images from two views.

TABLE I
MEAN REPROJECTION ERROR OF THE PROPOSED APPROACH (OBTAINED WITH THE SKULL), COMPARED WITH PREVIOUS STATE-OF-THE-ART APPROACHES

	Reprojection error (mm)	Tracking method	Connection
Cattari <i>et al.</i> [10]	2.02	RGB marker	Wired
Haxthausen <i>et al.</i> [9]	2.81	IR marker	Wired
Nguyen <i>et al.</i> [19]	1.19–3.02	ArUco markers	Wired
Proposed	0.983	SLAM	Wireless

SLAM module. The inclusion of an inertial measurement unit within the SLAM module enables accurate tracking, even when the controller is temporarily out of direct sight or under suboptimal lighting conditions. In contrast, visual markers require a constant line of sight, which is not always feasible in clinical environments.

IV. CONCLUSION

This work is the first to demonstrate wireless and marker-free volumetric ultrasound system, visualized and registered by an AR headset. By leveraging the robust tracking capabilities of the Magic Leap 2 controller, the wireless data transmission of the Clarius probe, and our novel differentiable calibration method, this system achieves a mean reprojection error < 1 mm. This system holds promise for surgical procedures such as intraoperative assessment, surgical training, and biopsies.

ACKNOWLEDGMENT

This work is supported by grant number 2022-309513 from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation, and the National Institutes of Health (NIH) R01 EB032960.

REFERENCES

- [1] C. Chen and M. A. Pertijs, "Integrated transceivers for emerging medical ultrasound imaging devices: A review," *IEEE Open Journal of the Solid-State Circuits Society*, vol. 1, pp. 104–114, 2021.
- [2] F. von Haxthausen, S. Böttger, D. Wulff, J. Hagenah, V. García-Vázquez, and S. Ipsen, "Medical robotics for ultrasound imaging: current systems and future trends," *Current Robotics Reports*, vol. 2, pp. 55–71, 2021.
- [3] H. T. Şen, M. A. L. Bell, Y. Zhang, K. Ding, E. Boctor, J. Wong, I. Iordachita, and P. Kazanzides, "System integration and in vivo testing of a robot for ultrasound guidance and monitoring during radiotherapy," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1608–1618, 2017.
- [4] S. Kim, H. J. Kang, A. Cheng, M. A. L. Bell, E. Boctor, and P. Kazanzides, "Photoacoustic image guidance for robot-assisted skull base surgery," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 592–597.
- [5] C. Peng, Q. Cai, M. Chen, and X. Jiang, "Recent advances in tracking devices for biomedical ultrasound imaging applications," *Micromachines*, vol. 13, no. 11, p. 1855, 2022.
- [6] C. Rüger, M. A. Feufel, S. Moosburner, C. Özbek, J. Pratschke, and I. M. Sauer, "Ultrasound in augmented reality: a mixed-methods evaluation of head-mounted displays in image-guided interventions," *International Journal of Computer Assisted Radiology and Surgery*, vol. 15, pp. 1895–1905, 2020.
- [7] T. Nguyen, W. Plishker, A. Matisoff, K. Sharma, and R. Shekhar, "Holous: Augmented reality visualization of live ultrasound images using hololens for ultrasound-guided procedures," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 2, pp. 385–391, 2022.
- [8] N. Costa, L. Ferreira, A. R. de Araújo, B. Oliveira, H. R. Torres, P. Morais, V. Alves, and J. L. Vilaça, "Augmented reality-assisted ultrasound breast biopsy," *Sensors*, vol. 23, no. 4, p. 1838, 2023.
- [9] F. Von Haxthausen, R. Moreta-Martinez, A. Pose Díez de la Lastra, J. Pascau, and F. Ernst, "Ultrasound: in situ visualization of live ultrasound images using hololens 2," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 11, pp. 2081–2091, 2022.
- [10] N. Cattari, S. Condino, F. Cutolo, M. Ferrari, and V. Ferrari, "In situ visualization for 3d ultrasound-guided interventions with augmented reality headset," *Bioengineering*, vol. 8, no. 10, p. 131, 2021.
- [11] A. Sorriento, M. B. Porfido, S. Mazzoleni, G. Calvosa, M. Tenucci, G. Ciuti, and P. Dario, "Optical and electromagnetic tracking systems for biomedical applications: A critical review on potentialities and limitations," *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 212–232, 2019.
- [12] E. Azimi, Z. Niu, M. Stiber, N. Greene, R. Liu, C. Molina, J. Huang, C.-M. Huang, and P. Kazanzides, "An interactive mixed reality platform for bedside surgical procedures," in *Medical Image Computing and Computer Assisted Intervention*. Springer, 2020, pp. 65–75.
- [13] L. Ma, T. Huang, J. Wang, and H. Liao, "Visualization, registration and tracking techniques for augmented reality guided surgery: a review," *Physics in Medicine & Biology*, vol. 68, no. 4, p. 04TR02, 2023.
- [14] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi, and G. Fichtinger, "Plus: open-source toolkit for ultrasound-guided intervention systems," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2527–2537, 2014.
- [15] J. Sola, J. Deray, and D. Atchuthan, "A micro lie theory for state estimation in robotics," *arXiv preprint arXiv:1812.01537*, 2018.
- [16] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [17] J. Tokuda, G. S. Fischer, X. Papademetris, Z. Yaniv, L. Ibanez, P. Cheng, H. Liu, J. Blevins, J. Arata, A. J. Golby *et al.*, "Openigtlink: an open network protocol for image-guided therapy environment," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 5, no. 4, pp. 423–434, 2009.
- [18] M. T. Graham, J. Huang, F. X. Creighton, and M. A. L. Bell, "Simulations and human cadaver head studies to identify optimal acoustic receiver locations for minimally invasive photoacoustic-guided neurosurgery," *Photoacoustics*, vol. 19, p. 100183, 2020.
- [19] K. W. Ng, Y. Gao, M. S. Furqan, Z. Yeo, J. Lau, K. Y. Ngiam, and E. T. Khoo, "Holopocus: Portable mixed-reality 3d ultrasound tracking, reconstruction and overlay," in *International Workshop on Advances in Simplifying Medical Ultrasound*. Springer, 2023, pp. 111–120.