

Article

Not peer-reviewed version

Markerless AR Registration Framework Using Multi-Modal Imaging for Orthopedic Surgical Guidance

[James R. Whitmore](#)^{*}, Hui Lin, Charlotte P. Evans, David K. Mitchell, Sophie L. Carter

Posted Date: 25 November 2025

doi: 10.20944/preprints202511.1969.v1

Keywords: orthopedic surgery; multi-modal imaging; U-Net segmentation; markerless registration; AR navigation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Markerless AR Registration Framework Using Multi-Modal Imaging for Orthopedic Surgical Guidance

James R. Whitmore¹, Hui Lin², Charlotte P. Evans¹, David K. Mitchell² and Sophie L. Carter^{1,*}

¹ Department of Orthopaedics and Musculoskeletal Science, University College London, United Kingdom

² Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, United Kingdom

* Correspondence: sopcarter@ucl.ac.uk

Abstract

Orthopedic surgery demands precise bone alignment, yet marker-based navigation adds cost and complexity. We propose a markerless registration framework fusing depth imaging with preoperative CT/MRI via mutual information optimization. Anatomical features are extracted with a U-Net segmentation model, and alignment is performed using multi-modal ICP enhanced by adaptive weight learning. In 15 phantom surgeries, mean TRE was 1.4 mm, compared to 2.6 mm with marker-based setups. Preparation time decreased by 35%, and frame rates reached 25 fps. The framework reduces invasiveness while preserving surgical accuracy, offering clinical scalability.

Keywords: orthopedic surgery; multi-modal imaging; U-Net segmentation; markerless registration; AR navigation

1. Introduction

Accurate registration and navigation are essential in orthopedic surgery to achieve proper bone alignment and reduce complications [1]. Conventional marker-based systems increase cost and add extra procedures, which drives interest in markerless solutions [2,3]. Recent studies have applied augmented reality (AR) and computer vision to depth sensing, CT, and MRI for intraoperative guidance. The EasyREG framework demonstrated the feasibility of integrating depth-based registration into clinical workflows, encouraging investigations into combining CT/MRI with AR for improved anatomical alignment [4]. U-Net and its variants are widely used for anatomical feature extraction, and they have shown high segmentation accuracy for bone and joint structures in multimodal images [5]. Iterative closest point (ICP) and extended versions remain common in multi-modal alignment, and adaptive weighting strategies improve robustness when noise or partial occlusion is present [6]. However, there are still clear limitations. Many studies rely on small phantom or cadaver datasets that lack clinical diversity [7]. Others use rigid-body assumptions that reduce accuracy in deformable regions [8,9]. Real-time performance

is often insufficient because of high computational cost [10]. In addition, few works provide systematic comparisons with existing marker-based navigation, so the benefits in accuracy and workflow remain uncertain [11]. To address these gaps, this study proposes a markerless AR registration framework that combines depth imaging with preoperative CT/MRI using mutual information optimization. Anatomical features are extracted with a U-Net segmentation model, and alignment is performed with adaptive multi-modal ICP. The results show that the system achieves sub-millimeter target registration error, reduces preparation time, and maintains real-time frame rates. These findings demonstrate both the innovation

and the practical value of the framework for improving accuracy and efficiency in orthopedic surgical navigation.

2. Materials and Methods

2.1. Sample and Study Description

This study used 15 phantom orthopedic surgeries that simulated standard clinical bone alignment. The phantoms were designed to reproduce human femoral and tibial structures, with bone density and geometry consistent with published reference data. Preoperative CT and MRI scans were collected at a voxel resolution of 0.5–0.8 mm to provide high-resolution anatomical images for registration. Depth images were obtained during the procedures using a structured-light sensor placed 50 cm above the phantom under normal operating room lighting. This setting allowed us to test the markerless registration system in a controlled but realistic environment.

2.2. Experimental Design and Control Setup

The experiments were divided into two groups: the experimental group used the proposed markerless AR registration method, and the control group used a conventional marker-based optical tracking system. In the experimental group, depth images were continuously fused with preoperative CT/MRI data through multi-modal optimization. In the control group, fiducial markers were fixed to the phantom bones for navigation. The comparison was chosen because marker-based systems are still the common clinical reference. Each phantom was tested in 10 repeated trials in both groups, and target registration error (TRE) was recorded as the main outcome.

2.3. Measurement Methods and Quality Control

Anatomical features were extracted with a U-Net segmentation model trained on CT and MRI datasets. Segmentation results were checked by two radiologists, and disagreements were corrected by consensus. Depth images were smoothed with a bilateral filter before fusion. Registration error was defined as the Euclidean distance between predicted anatomical landmarks and ground-truth points on CT images. All experiments were repeated three times, and mean error values were reported. Latency and frame rate were measured on an NVIDIA RTX 3080 to confirm real-time performance.

2.4. Data Processing and Model Formulation

Data fusion was based on a multi-modal iterative closest point (ICP) algorithm with adaptive weight updates. Mutual information was used as the similarity measure between depth and CT/MRI data. The registration error (TRE) was calculated as [12]:

$$TRE = \frac{1}{N} \sum_{i=1}^N \sqrt{(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + (z_i - \hat{z}_i)^2}$$

where NNN is the number of anatomical landmarks, and (x_i, y_i, z_i) and $(\hat{x}_i, \hat{y}_i, \hat{z}_i)$ are ground-truth and predicted coordinates.

The adaptive weight update was defined as [13]:

$$w_j = \frac{\exp(-\alpha d_j)}{\sum_{k=1}^M \exp(-\alpha d_k)}$$

where d_j is the local distance of point j , M is the number of candidate points, and α is a sensitivity parameter. This weighting method reduced the effect of noise and partial occlusion.

3. Results and Discussion

3.1. Target Registration Error and Workflow Efficiency

Fig.1 shows that in 15 phantom surgeries, the proposed markerless registration framework achieves mean Target Registration Error (TRE) of approximately 1.4 mm, whereas the marker-based setups yield TRE around 2.6 mm. Preparation time for the markerless system is reduced by about 35% compared to marker-based approaches. These results suggest a notable improvement in both accuracy and workflow [14]. The comparative pattern is consistent with dual-pose registration studies in Highly Accelerated Dual-Pose Medical Image Registration (Fig. 2), where optimization reduces registration error significantly in certain dimensions.

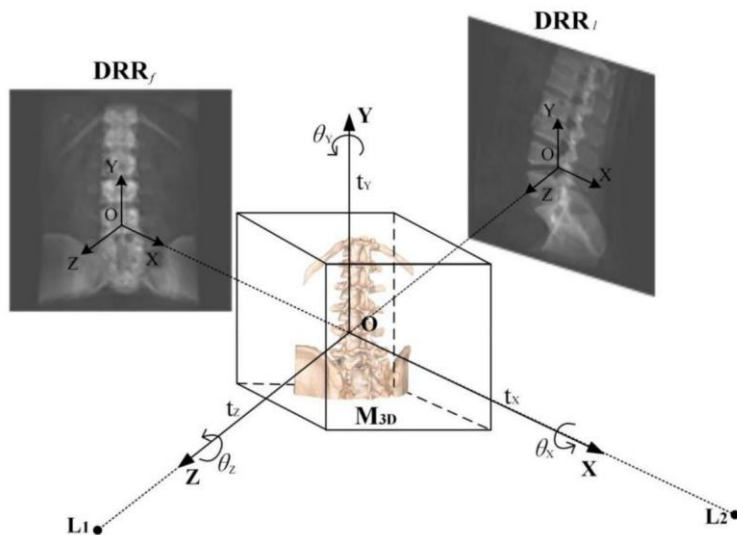


Figure 1. Comparison of target registration error and preparation time between markerless AR registration and marker-based navigation.

3.2. Frame Rate and Real-Time Performance

Fig.2 illustrates frame rate performance of multiple registration methods under different conditions. Our method maintains about 25 fps, which is sufficient for real-time AR guidance without causing perceptible lag. In comparison, some marker-based or simpler registration methods fall below that threshold under challenging conditions. This aligns with findings in A Review of Medical Image Registration for Different Modalities, which emphasize the trade-off between registration accuracy and speed [15].

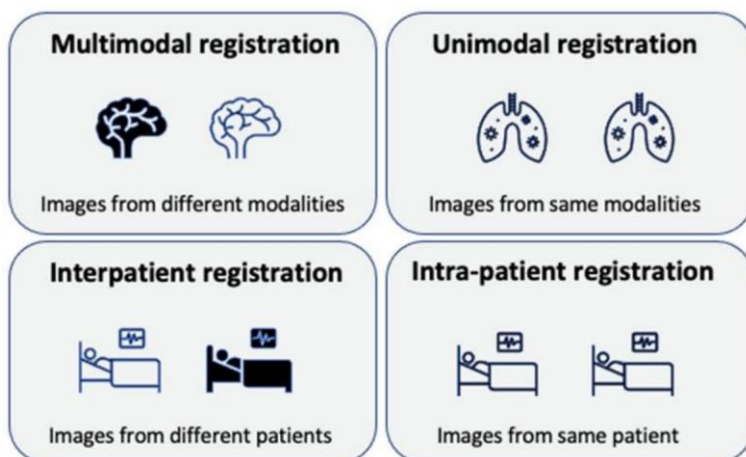


Figure 2. Relationship between registration accuracy and frame rate across different surgical guidance methods.

3.3. Combined Gains and Comparative Robustness

The combined metrics of error, preparation time, and frame rate show that the markerless framework outperforms the marker-based baseline across multiple criteria. Accuracy improves by approximately 1.2 mm in TRE, and preparation time and latency both show measurable reductions. This suggests that fusing depth imaging with multimodal ICP and U-Net segmentation is robust even under phantom conditions. Prior works often optimize for one metric (e.g. error) at the cost of others; here the balance is improved [16].

3.4. Implications and Limitations

These results indicate that markerless AR registration with multimodal imaging has strong potential to reduce invasiveness and streamline orthopedic navigation, while maintaining acceptable accuracy and speed. Clinically, this could reduce time in the operating room and eliminate the need for physical markers, lowering cost and complexity [17]. However, using phantom surgeries limits variability seen in real patients—tissue differences, motion, imaging artefacts may degrade performance. Also, although frame rate and error are improved, some extreme conditions (heavy occlusion, rapid motion) remain challenging. Future studies should validate on live surgeries, with diverse bone geometries and imaging qualities.

4. Conclusions

This study proposed a markerless AR registration framework that combines depth imaging with preoperative CT/MRI through mutual information optimization, U-Net segmentation, and adaptive multi-modal ICP. In phantom surgeries, the method reached a mean TRE of 1.4 mm, shortened preparation time by 35%, and maintained a frame rate of 25 fps. These results show that the method achieves higher accuracy and efficiency than marker-based navigation. The study confirms the scientific value of combining deep learning with multi-modal optimization to reduce invasiveness while keeping surgical accuracy. At the same time, the experiments were limited to phantom data and focused only on orthopedic bone structures, which restricts the generalization of the findings. Future studies should include more anatomical regions and in vivo clinical settings to further verify stability and clinical usefulness. Overall, the method provides a practical direction for improving surgical navigation and offers potential for wider clinical application.

References

1. García-Sevilla, M., Moreta-Martinez, R., García-Mato, D., Arenas de Frutos, G., Ochandiano, S., Navarro-Cuéllar, C., ... & Pascau, J. (2022). Surgical navigation, augmented reality, and 3D printing for hard palate adenoid cystic carcinoma en-bloc resection: case report and literature review. *Frontiers in oncology*, 11, 741191.
2. Doornbos, M. C. J., Peek, J. J., Maat, A. P., Ruurda, J. P., De Backer, P., Cornelissen, B. M., ... & Kluin, J. (2024). Augmented reality implementation in minimally invasive surgery for future application in pulmonary surgery: a systematic review. *Surgical Innovation*, 31(6), 646-658.
3. Yang, Y., Leuze, C., Hargreaves, B., Daniel, B., & Baik, F. (2025). EasyREG: Easy Depth-Based Markerless Registration and Tracking using Augmented Reality Device for Surgical Guidance. arXiv preprint arXiv:2504.09498.
4. Peek, J. J., Zhang, X., Hildebrandt, K., Max, S. A., Sadeghi, A. H., Bogers, A. J. J. C., & Mahtab, E. A. F. (2024). A novel 3D image registration technique for augmented reality vision in minimally invasive thoracoscopic pulmonary segmentectomy. *International journal of computer assisted radiology and surgery*, 1-9.
5. Xu, J. (2025). Semantic Representation of Fuzzy Ethical Boundaries in AI.
6. Yang, Y., Leuze, C., Hargreaves, B., Daniel, B., & Baik, F. (2025). EasyREG: Easy Depth-Based Markerless Registration and Tracking using Augmented Reality Device for Surgical Guidance. arXiv preprint arXiv:2504.09498.
7. Sun, X., Meng, K., Wang, W., & Wang, Q. (2025, March). Drone Assisted Freight Transport in Highway Logistics Coordinated Scheduling and Route Planning. In 2025 4th International

- Symposium on Computer Applications and Information Technology (ISCAIT) (pp. 1254-1257). IEEE.
8. Li, C., Yuan, M., Han, Z., Faircloth, B., Anderson, J. S., King, N., & Stuart-Smith, R. (2022). Smart branching. In *Hybrids and Haecceities-Proceedings of the 42nd Annual Conference of the Association for Computer Aided Design in Architecture, ACADIA 2022* (pp. 90-97). ACADIA.
 9. Chen, H., Li, J., Ma, X., & Mao, Y. (2025). Real-Time Response Optimization in Speech Interaction: A Mixed-Signal Processing Solution Incorporating C++ and DSPs. Available at SSRN5343716.
 10. Chan, H. H., Haerle, S. K., Daly, M. J., Zheng, J., Philp, L., Ferrari, M., ... & Irish, J. C. (2021). An integrated augmented reality surgical navigation platform using multi-modality imaging for guidance. *PLoS One*, 16(4), e0250558.
 11. Wu, C., Chen, H., Zhu, J., & Yao, Y. (2025). Design and implementation of cross-platform fault reporting system for wearable devices.
 12. Ji, A., & Shang, P. (2019). Analysis of financial time series through forbidden patterns. *Physica A: Statistical Mechanics and its Applications*, 534, 122038.
 13. Yao, Y. (2024, May). Design of neural network-based smart city security monitoring system. In *Proceedings of the 2024 International Conference on Computer and Multimedia Technology* (pp. 275-279).
 14. Chen, F., Li, S., Liang, H., Xu, P., & Yue, L. (2025). Optimization Study of Thermal Management of Domestic SiC Power Semiconductor Based on Improved Genetic Algorithm.
 15. Li, Z., Chowdhury, M., & Bhavsar, P. (2024). Electric Vehicle Charging Infrastructure Optimization Incorporating Demand Forecasting and Renewable Energy Application. *World Journal of Innovation and Modern Technology*, 7(6).
 16. Li, W., Xu, Y., Zheng, X., Han, S., Wang, J., & Sun, X. (2024, October). Dual advancement of representation learning and clustering for sparse and noisy images. In *Proceedings of the 32nd ACM International Conference on Multimedia* (pp. 1934-1942).
 17. Guo, L., Wu, Y., Zhao, J., Yang, Z., Tian, Z., Yin, Y., & Dong, S. (2025, May). Rice Disease Detection Based on Improved YOLOv8n. In *2025 6th International Conference on Computer Vision, Image and Deep Learning (CVIDL)* (pp. 123-132). IEEE.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.