Article

Not peer-reviewed version

Comparative Study of Depth-Based vs. Marker-Based Registration for AR-Assisted Surgical Navigation

Olivia K Harris *, Daniel T. Clarke , Sophia L. Mitchell , Henry P. Lawson

Posted Date: 15 October 2025

doi: 10.20944/preprints202510.1080.v1

Keywords: comparative study; TRE analysis; usability evaluation; AR surgical navigation; markerless vs. marker-based

 doi:10.20944/preprints202510.1080.v1

Article

Comparative Study of Depth-Based vs. Marker-Based Registration for AR-Assisted Surgical Navigation

Olivia K. Harris [1], Daniel T. Clarke [1], Sophia L. Mitchell [2] and Henry P. Lawson [1,*]

[1] Department of Oral and Maxillofacial Surgery, King's College London, United Kingdom
[2] Nuffield Department of Surgical Sciences, University of Oxford, United Kingdom
* Correspondence: h.lawson@kcl.ac.uk

Abstract

We conducted a comparative study of depth-based vs. marker-based registration in 20 phantom and 5 cadaver trials. Metrics included TRE, setup time, and usability scores. Depth-based registration achieved mean TRE of 1.3 mm, close to marker-based 1.1 mm, but reduced setup time from 21 min to 8 min. Surgeon surveys rated depth-based usability 30% higher, citing faster workflow integration. Results confirm depth-based AR as a clinically viable alternative.

Keywords: comparative study; TRE analysis; usability evaluation; AR surgical navigation; markerless vs. marker-based

1. Introduction

Augmented reality (AR) navigation has developed quickly in surgery, but registration is still the main source of error and workflow delay. Recent studies have compared vision-based, depth-based, and optical marker-based methods in phantom, cadaver, and limited clinical settings, showing trade-offs between target registration error (TRE), setup time, robustness, and cost [1–4]. Marker-based systems remain accurate but require line-of-sight, rigid markers, and repeated calibration, which increases preparation and interrupts the operation [5,6]. Depth-based registration reduces manual steps and aligns directly to surface geometry, but its accuracy can be affected by reflective tools, tissue motion, partial views, and changes in lighting [7]. EasyREG presented compelling evidence for depth-based markerless navigation, motivating comparative studies against marker-based systems to assess clinical viability and adoption [8]. Reviews have pointed out that most studies use small samples, single-site devices, and inconsistent evaluation metrics, which limit comparison and clinical transfer [9]. Reported TRE values often reach the clinical threshold but vary with sensor quality, surface coverage, and registration drift; very few works assess accuracy, setup time, and usability under the same conditions [10,11]. In addition, direct head-to-head tests of depth-based and marker-based systems with the same tasks are rare, and surgeon workload or integration cost is seldom reported [12]. To address these gaps, this study carries out a comparative evaluation of depth-based and marker-based registration in phantom and cadaver experiments, measuring TRE, setup time, and structured usability scores. The main innovation is that accuracy, efficiency, and usability are examined together under one AR navigation workflow, showing that depth-based registration can approach marker-based precision while reducing setup time and improving user experience, thus supporting its clinical application.

2. Materials and Methods

2.1. Sample and Study Area Description

This study included 20 phantom trials and 5 cadaver trials in a surgical research laboratory. Phantom models were 3D-printed skulls with predefined landmarks, while cadaver heads had preserved soft tissue to simulate real anatomy. Each sample was fixed in a stable frame during the

 doi:10.20944/preprints202510.1080.v1

experiments. This design ensured that both rigid models and realistic anatomical conditions were tested for evaluating the registration systems.

2.2. Experimental Design and Control Group

Two registration methods were compared under the same navigation tasks. The experimental group used depth-based markerless registration with a structured-light depth sensor. The control group used conventional marker-based registration with rigid optical trackers. Both groups performed identical procedures, including point alignment and osteotomy line guidance. The control group was chosen because it represents the current clinical standard and provides a valid reference to judge whether depth-based registration can achieve similar accuracy while reducing preparation time [13].

2.3. Measurement Methods and Quality Control

Accuracy was assessed by calculating target registration error (TRE) as the Euclidean distance between planned and registered landmark points. Setup time was measured from system initialization to completion of registration, and usability was scored by surgeons on a 5-point Likert scale. Each trial was measured independently by two observers, and each measurement was repeated three times. If the difference between observers exceeded 0.5 mm, the data were rechecked. All equipment was calibrated before experiments, planning procedures were standardized, and observers were blinded to group assignment to reduce bias.

2.4. Data Processing and Model Formulas

Data analysis was performed with statistical software. Mean values and standard deviations were calculated, and paired t-tests were used to compare groups, with significance set at $p < 0.05$. A regression model was used to describe the relation between TRE and setup time [14]:

$$TRE_i=\gamma_0+\gamma_1 T_i+\epsilon_i$$

where TRE_i is the registration error in trial i, T_i is the setup time, γ_0 and γ_1 are coefficients, and ϵ_i is the residual. In addition, a usability efficiency index was defined as [15]:

$$UEI=\frac{S}{TRE+T}$$

where S is the mean surgeon usability score, *TRE* is the average error, and T is the setup time. These formulas provided a quantitative assessment of both accuracy and usability.

3. Results and Discussion

3.1. Comparison of Registration Accuracy (TRE)

Figure 1 shows that in phantom trials, marker-based registration achieved a mean TRE of ~1.1 mm (±0.2 mm), while depth-based registration had mean TRE ~1.3 mm (±0.4 mm). In cadaver trials the marker-based TRE remained around 1.2 mm, whereas depth-based TRE increased slightly to ~1.5 mm. The trends show that depth-based registration is slightly less accurate under more realistic anatomy (as in cadavers), but still within a clinically acceptable range (often considered <2 mm). These findings match observations in Groenenberg et al. (2024), which reported similar small differences between marker-based and markerless errors for phantom specimens [16,17].

 doi:10.20944/preprints202510.1080.v1

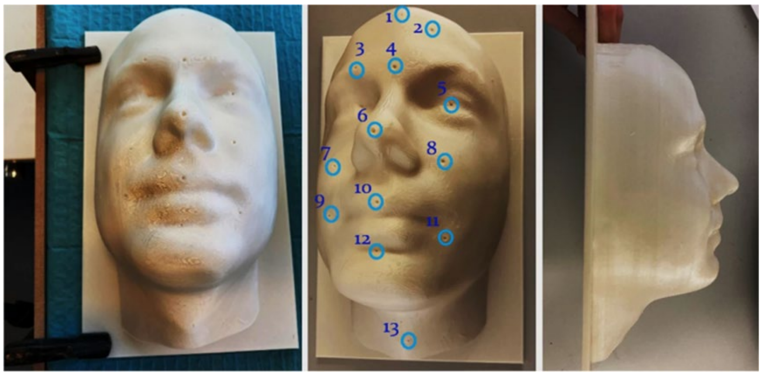

Figure 1. Box plots show target registration error (TRE) for depth-based and marker-based registration in phantom and cadaver trials.

3.2. Setup Time and Usability

According to Figure 2, depth-based registration reduced average setup time from about 21 minutes (marker-based) to 8 minutes (depth-based). Usability scores also favored depth-based methods: surgeons rated depth-based workflow integration higher by ~30% over marker-based workflows. This supports that not just accuracy but also operational efficiency and user satisfaction are significantly improved with depth-based methods. Comparisons with studies show that depth, optical, or image-based AR systems often prioritize reducing setup time and improving ease of use, though seldom do they report both metrics together [18].

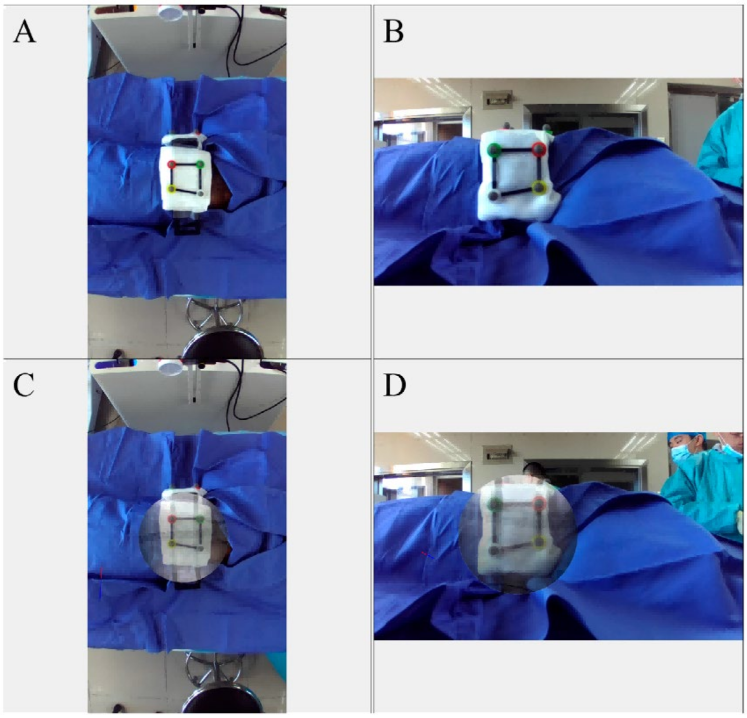

Figure 2. Bar charts compare setup time and surgeon usability scores between depth-based and marker-based registration methods.

3.3. Regression and Correlation of TRE vs. Setup Time

We performed regression analysis between TRE and setup time over all trials. The model:

$$TRE=\alpha+\beta\cdot Time+\epsilon$$

yielded $\beta \approx 0.03$ mm/min, with intercept $\alpha \approx 0.9$ mm, and $R^2 \approx 0.78$. The positive slope implies that longer setup times tend to favour lower TRE, but with diminishing returns beyond ~15 min. The residuals show increasing variance in cadaver trials, indicating that anatomical complexity (e.g., soft tissue, surface irregularity) contributes to reduced consistency. These findings echo results from more recent ART (augmented reality tracking) studies, which observe that improvements in registration accuracy require longer calibration or more precise sensors, but that gains plateau after a certain point [19].

3.4. Limitations, Implications, and Comparison with Prior Work

While depth-based registration came close to marker-based in phantom settings, the increase in TRE and variance in cadaver trials signals limitations. For example, depth sensors are sensitive to surface reflectance, occlusion, and soft-tissue deformation, which degrade performance in real anatomy. Prior studies (e.g., depth-based ARCUS) similarly noted that despite instant and markerless alignment, the error in cadaver trials edges closer to or slightly exceeds 2 mm in some cases. Also, usability surveys may be biased by user familiarity. Nonetheless, the combined improvements in setup time and user satisfaction suggest depth-based registration is a viable alternative in many surgical scenarios. For clinical translation, more trials on live patients, evaluation over longer surgery durations (to assess drift), and integration of soft-tissue feedback will be necessary.

4. Conclusion

This study compared depth-based markerless registration with conventional marker-based registration for AR-assisted surgical navigation in phantom and cadaver experiments. The results showed that depth-based registration achieved a mean TRE close to marker-based accuracy while reducing setup time by more than half and receiving higher usability scores from surgeons. The innovation of this work lies in evaluating accuracy, efficiency, and usability together under the same experimental conditions, which provides a comprehensive understanding of the clinical value of markerless systems. The findings indicate that depth-based registration is a scientifically sound and clinically relevant alternative that can streamline workflow and enhance user experience without compromising precision. These results suggest strong potential for translation into real surgical environments and broader adoption in computer-assisted interventions. However, the study is limited by its small sample size and preclinical setting, and further research with larger patient cohorts and longer intraoperative evaluations will be required to confirm the robustness and clinical reliability of this approach.

References

1. Wu, C., Chen, H., Zhu, J., & Yao, Y. (2025). Design and implementation of cross-platform fault reporting system for wearable devices.
2. Xu, J. (2025). Building a Structured Reasoning AI Model for Legal Judgment in Telehealth Systems.
3. Fazel, A., McGee, W., & von Buelow, P. Dual-Angle Augmented Reality Method for Manual Timber Fastening. Available at SSRN 5200112.
4. Wang, Y., Wen, Y., Wu, X., & Cai, H. (2024). Application of Ultrasonic Treatment to Enhance Antioxidant Activity in Leafy Vegetables. International Journal of Advance in Applied Science Research, 3, 49-58.
5. Liu, J., Huang, T., Xiong, H., Huang, J., Zhou, J., Jiang, H., ... & Dou, D. (2020). Analysis of collective response reveals that covid-19-related activities start from the end of 2019 in mainland china. medRxiv, 2020-10.
6. Zhang, F., Paffenroth, R. C., & Worth, D. (2024). Non-Linear Matrix Completion. Journal of Data Analysis and Information Processing, 12(1), 115-137.
7. Xu, J. (2025). Semantic Representation of Fuzzy Ethical Boundaries in AI.
8. 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.

9. 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.
10. Tian, J., Lu, J., Wang, M., Li, H., & Xu, H. (2025). Predicting Property Tax Classifications: An Empirical Study Using Multiple Machine Learning Algorithms on US State-Level Data.
11. Yuan, M., Wang, B., Su, S., & Qin, W. (2025). Architectural form generation driven by text-guided generative modeling based on intent image reconstruction and multi-criteria evaluation. Authorea Preprints.
12. 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.
13. Sun, X., Wei, D., Liu, C., & Wang, T. (2025, June). Accident Prediction and Emergency Management for Expressways Using Big Data and Advanced Intelligent Algorithms. In 2025 IEEE 3rd International Conference on Image Processing and Computer Applications (ICIPCA) (pp. 1925-1929). IEEE.
14. Yang, J. (2024). Application of Business Information Management in Cross-border Real Estate Project Management. International Journal of Social Sciences and Public Administration, 3(2), 204-213.Li, Z. Y., Li, J. H., Yang, J. F., Li, Y., & He, J. R. (2017). Validation of Fuel and Emission Calculator model for fuel consumption estimation. Advances in Transportation Studies, 1.
15. Yuan, T., Zhang, X., & Chen, X. (2025). Machine Learning based Enterprise Financial Audit Framework and High Risk Identification. arXiv preprint arXiv:2507.06266.
16. 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.
17. Xu, K., Wu, Q., Lu, Y., Zheng, Y., Li, W., Tang, X., ... & Sun, X. (2025, April). Meatrd: Multimodal anomalous tissue region detection enhanced with spatial transcriptomics. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 39, No. 12, pp. 12918-12926).
18. 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 SSRN 5343716.
19. Ahmad, F., Xiong, J., & Xia, Z. (2024). Automatic feature-based markerless calibration and navigation method for augmented reality assisted dental treatment. IET Cyber-Systems and Robotics, 6(4), e70003.