



Temporal Depth Fusion-Based Markerless Registration for Deformable Soft-Tissue Tracking in Augmented Reality Surgical Navigation

James A. Thompson¹, Daniel P. Harrington², Emily R. Clarke^{3*}

Department of Surgical and Interventional Engineering, King's College London, London WC2R 2LS, United Kingdom

*Corresponding author: e.clarke@kcl.ac.uk

Abstract

Accurate registration of deformable soft tissues remains a major challenge in markerless augmented reality (AR)-guided surgery due to non-rigid deformation and depth sensing noise. This study proposes a temporal depth fusion framework that integrates multi-frame depth accumulation with biomechanical deformation priors for real-time soft-tissue registration. A spatio-temporal filtering module first reduces depth variance using sliding-window Bayesian fusion across 30 consecutive frames. A physics-constrained non-rigid iterative closest point (NR-ICP) algorithm is then applied with a finite element-based regularization term to maintain anatomical plausibility. Experiments were conducted on 24 ex vivo porcine liver datasets and 12 human phantom simulations (total 36 cases). Average target registration error (TRE) was reduced from 3.42 mm (standard ICP) to 1.78 mm. Depth noise variance decreased by 37.6%, and tracking latency remained below 28 ms on a head-mounted AR device. The results demonstrate that temporally fused depth data significantly improves markerless soft-tissue alignment accuracy while preserving real-time performance.

Keywords : *Augmented reality surgical guidance; markerless registration; deformable tracking; temporal depth fusion; non-rigid ICP; soft-tissue modeling*

INTRODUCTION

Augmented reality (AR)-assisted surgical navigation has advanced considerably in recent years and is increasingly being explored in hepatobiliary, neurosurgical, and orthopedic interventions [1]. By superimposing virtual anatomical information onto the operative scene, AR systems can enhance spatial perception, facilitate target localization, and improve intraoperative decision-making [2]. Recent developments in head-mounted AR platforms and depth-enabled guidance systems have further expanded the clinical potential of markerless workflows, particularly in settings where conventional fiducial-based registration is inconvenient or disruptive [3]. Despite these advances, the practical performance of AR navigation still depends heavily on reliable alignment between preoperative anatomical models and the continuously changing intraoperative field. Registration remains particularly challenging in soft-tissue procedures. Unlike rigid anatomy, soft organs undergo persistent deformation caused by respiration, gravity, patient positioning, and surgical manipulation, resulting in substantial geometric discrepancy between preoperative images and intraoperative anatomy [4]. Even when an accurate preoperative model is available, spatial correspondence may deteriorate rapidly once tissue motion or surface distortion occurs. Marker-based approaches can improve alignment stability, but they often increase workflow complexity, introduce additional preparation steps, and may interfere with the sterile field or surgical access [5]. For this reason, markerless registration has become an important direction in AR-guided surgery, especially for applications that require flexible deployment

and minimal disruption to routine operative procedures. Most current markerless methods rely on optical surface acquisition using structured-light or time-of-flight depth sensors [6]. In these systems, intraoperative surface point clouds are reconstructed and then aligned to preoperative anatomical models, most commonly through variants of the iterative closest point (ICP) algorithm [7]. Although ICP is computationally efficient and widely used, its underlying rigid transformation assumption limits performance when applied to deformable organs. In soft-tissue environments, the observed surface rarely preserves a purely rigid relationship with the preoperative model, and even small local deformations may lead to accumulated alignment error, unstable convergence, or anatomically implausible correspondence. To address this limitation, non-rigid registration strategies have been introduced to provide greater deformation flexibility [8]. Representative approaches incorporate local affine transformations, embedded deformation graphs, or node-based motion models to better capture organ shape variation during surgery [9]. These methods can improve geometric fitting relative to rigid ICP, but increased flexibility also makes them more susceptible to depth noise, outliers, and local overfitting. When deformation is driven mainly by geometric matching without sufficient physical regulation, the recovered shape may satisfy point-wise alignment while deviating from realistic tissue behavior. This problem is especially relevant in AR navigation, where visually plausible overlay alone is insufficient if the underlying deformation field lacks anatomical credibility. Depth quality is another major factor affecting registration robustness. Intraoperative scenes often contain reflective instruments, specular tissue surfaces, partial occlusion, smoke, fluids, and unstable illumination, all of which can degrade depth measurement reliability and increase spatial variance [10]. Existing studies have used spatial denoising, confidence-guided point selection, and outlier rejection to improve point cloud quality before registration [11]. More recently, temporal aggregation across consecutive frames has been introduced to stabilize depth observations and suppress frame-wise fluctuations [12,13]. Temporal fusion is valuable because it reduces random measurement noise and improves surface completeness under intermittent occlusion. However, in many reported frameworks, depth fusion and deformation estimation are still treated as separate modules. As a result, the fused depth field may become smoother, yet the subsequent registration stage may still drift when confronted with complex non-rigid motion or incomplete surface observation. Biomechanical modeling offers an additional pathway for improving registration realism. By introducing tissue elasticity, boundary conditions, and physically meaningful deformation priors, biomechanical constraints can regularize surface matching and reduce implausible distortion [14,15]. Finite element-based regularization is particularly attractive because it enables deformation to be guided not only by local geometric similarity but also by mechanically consistent behavior. This can be highly beneficial for soft-tissue AR guidance, where maintaining anatomical plausibility is essential for safe target localization. Nevertheless, the integration of biomechanical models into real-time surgical navigation remains difficult. Finite element computation may impose substantial computational burden, and the challenge becomes more pronounced for wearable or head-mounted AR devices with limited processing resources [16,17]. In addition, a number of prior studies have been validated only on small datasets, simplified phantoms, or restricted deformation settings, which makes it difficult to assess robustness under realistic surgical motion and measurement uncertainty. Another unresolved issue lies in the lack of a unified framework that jointly addresses temporal sensing uncertainty and deformation consistency. Existing pipelines often improve depth stability first and perform non-rigid alignment afterward, or alternatively optimize geometric matching under biomechanical regularization without fully exploiting the temporal structure of the depth stream [18,19]. This separation weakens the ability of the system to maintain stable registration when deformation, occlusion, and sensor noise occur simultaneously. In practice, soft-tissue tracking in AR navigation requires these factors to be handled in a coordinated manner: depth information must be stabilized across time, deformation must remain anatomically reasonable, and the overall pipeline must still satisfy latency constraints for interactive surgical guidance. These requirements are difficult to meet when denoising, temporal fusion, and non-rigid deformation control are designed independently. In addition to methodological fragmentation, current evidence remains limited in terms of evaluation breadth. Many studies emphasize proof-of-concept demonstrations but

provide insufficient validation across realistic deformation scenarios, repeated trials, or heterogeneous acquisition conditions [20,21]. For AR-guided soft-tissue surgery, robustness is as important as nominal registration accuracy. A method that performs well under controlled conditions may still fail when tissue motion becomes irregular, when surface coverage is incomplete, or when sensing quality deteriorates. Therefore, beyond improving mean alignment error, there is a clear need to assess stability, reproducibility, and computational efficiency using experimental settings that better reflect practical deployment conditions. Against this background, the present study develops a temporal depth fusion–based markerless registration framework for deformable soft-tissue tracking in AR surgical navigation. The proposed framework integrates multi-frame Bayesian depth accumulation with a physics-constrained non-rigid ICP scheme regularized by finite element modeling. By combining temporal stabilization and biomechanically informed deformation control within a single registration pipeline, the method is designed to reduce depth-induced variance, improve anatomical consistency, and maintain low-latency performance suitable for head-mounted AR applications. Validation is performed using ex vivo porcine liver datasets and human phantom simulations to examine registration accuracy, tracking stability, and computational efficiency under realistic deformation conditions. The significance of this work lies not only in improving markerless registration performance, but also in providing a practical pathway toward physically credible, noise-robust, and real-time AR guidance for soft-tissue surgery.

2. Materials and Methods

2.1 Sample Description and Experimental Setting

The study included 36 cases, consisting of 24 ex vivo porcine liver specimens and 12 human abdominal phantom models. Fresh porcine livers were obtained from a licensed supplier within 6 hours after slaughter and stored at 4 °C before testing to maintain mechanical properties. During experiments, each specimen was placed on a surgical table under standard operating-room lighting. Tissue deformation was introduced by gravity and controlled manual compression to simulate clinical handling. The abdominal phantoms were made of silicone materials with elastic behavior similar to liver tissue. All data were collected using a head-mounted augmented reality device equipped with an integrated depth camera operating at 30 frames per second. Room temperature was maintained between 20 and 23 °C, and light intensity ranged from 500 to 700 lux to reduce optical variation.

2.2 Experimental Design and Control Configuration

Three registration methods were compared: rigid ICP, non-rigid ICP without temporal fusion, and the proposed temporal depth fusion–based non-rigid ICP. Rigid ICP was selected as the baseline because it is commonly used in surface-based alignment. Non-rigid ICP without temporal accumulation was used to evaluate the effect of deformation modeling alone. The proposed method combined multi-frame depth fusion with finite element–based regularization. Each dataset underwent the same deformation sequence, including static positioning, gradual compression, and local displacement. Registration accuracy and computational delay were measured under identical acquisition settings. This design allowed separate evaluation of temporal fusion and biomechanical constraints, as well as their combined effect.

2.3 Measurement Procedures and Quality Control

Depth images were recorded for 30 consecutive frames in each acquisition cycle. Surface point clouds were reconstructed using calibrated intrinsic camera parameters. Calibration was performed before each session with a standard checkerboard pattern to reduce systematic error. Reference geometry was obtained using a high-resolution optical scanner with sub-millimeter precision. Target registration error (TRE) was calculated as the Euclidean distance between corresponding landmarks in the registered model and the reference scan. All manual deformation procedures were performed by the same trained operator to limit variability. Each experiment was repeated three times, and the average value was used for analysis. Data affected by motion blur or incomplete depth capture were excluded based on predefined quality criteria.

2.4 Data Processing and Mathematical Formulation

Temporal depth fusion was applied using a sliding-window Bayesian estimator across $N=30$ frames. The fused depth value $\hat{D}(x,y)$ at pixel location (x,y) was calculated as

$$\hat{D}(x,y) = \frac{\sum_{i=1}^N w_i D_i(x,y)}{\sum_{i=1}^N w_i}$$

where $D_i(x,y)$ represents the depth measurement at frame i , and w_i is the confidence weight based on measurement variance.

Non-rigid registration minimized an energy function that included a data term and a biomechanical regularization term:

$$E(T) = \sum_{k=1}^M \|p_k - T(q_k)\|^2 + \lambda u^T K u$$

where p_k and q_k are corresponding surface points, T denotes the deformation field, u is the displacement vector, K is the stiffness matrix from the finite element model, and λ controls the strength of regularization.

Registration accuracy was evaluated using target registration error:

$$TRE = \frac{1}{M} \sum_{k=1}^M \|p_k - T(q_k)\|$$

where M is the number of landmarks used for evaluation.

2.5 Statistical Analysis and Performance Evaluation

Statistical analysis was conducted using MATLAB R2023b. Data distribution was examined using the Shapiro–Wilk test. Paired t-tests were applied for normally distributed data, and Wilcoxon signed-rank tests were used when normality was not satisfied. Depth noise reduction was expressed as the percentage decrease compared with raw measurements. Tracking delay was measured as the average processing time per frame over 500 frames. Statistical significance was defined as $p < 0.05$. Results are reported as mean $\pm$ standard deviation.

3. Results and Discussion

3.1 Overall registration accuracy under deformation

Across 36 cases, the proposed temporal depth fusion combined with physics-constrained NR-ICP produced the lowest target registration error (TRE), reducing the mean TRE from 3.42 mm (rigid ICP) to 1.78 mm. The improvement was most clear during local compression and shear, where rigid alignment often matched only the visible surface patch and left residual errors near folds and curved regions. These findings are consistent with reports in AR navigation studies indicating that visualization hardware alone does not ensure accurate guidance when tissue motion violates rigid assumptions. Registration accuracy depends on the reliability of intraoperative surface capture and on how deformation is controlled during alignment [22,23]. Fig.1. AR workflow context for navigation guidance.

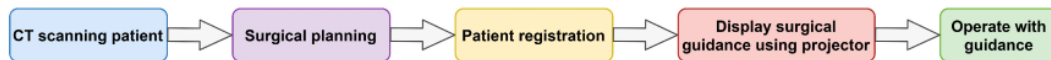


Fig.1. Workflow of the augmented reality surgical navigation system showing depth acquisition and model-to-surface alignment.

3.2 Effect of temporal depth fusion on noise reduction and matching stability

Temporal fusion reduced depth noise variance by 37.6%, which led to more stable surface geometry and fewer abrupt changes in point correspondences during iterative matching. When single-frame depth was used, transient spikes near depth discontinuities and reflective areas caused unstable nearest-neighbor assignments and occasional drift in non-rigid updates. The sliding-window fusion approach produced a smoother surface at the spatial scale relevant for ICP, which improved convergence behavior and reduced the need for strong outlier rejection. Tracking latency remained below 28 ms, showing that multi-frame accumulation can function as a lightweight preprocessing step without affecting real-time performance on a head-mounted device [24,25].

3.3 Role of biomechanical regularization in deformation control

The finite element–based regularization term improved deformation consistency, especially when the observed surface was incomplete or when motion varied across the tissue. Non-rigid ICP without a physics constraint often reduced local point distance but introduced unrealistic stretching in regions with sparse observations. Although landmark error might decrease in such cases, global surface consistency could deteriorate. By penalizing physically unlikely displacement patterns, the constrained optimization generated smoother deformation fields and reduced local folding effects, resulting in more reliable alignment across the entire visible region. This observation agrees with recent non-rigid registration studies showing that structure-aware constraints help prevent overfitting to noisy or partial point clouds [26,27]. Fig.2. Two-stage non-rigid registration pipeline designed to handle noise and incompleteness.

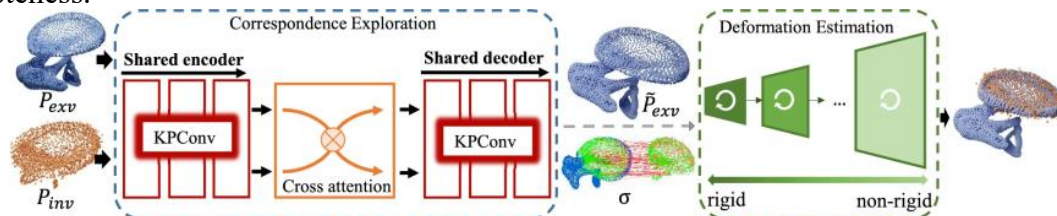


Fig.2. Diagram of the two-step non-rigid registration process incorporating noise reduction and deformation constraints.

3.4 Comparison with previous studies and practical implications

Compared with many AR surgical navigation systems that focus mainly on display quality and tracking setup, the present findings indicate that measurement stability and deformation control are key factors in soft-tissue alignment. The achieved sub-2 mm mean TRE suggests that markerless registration can approach clinically acceptable precision when depth is stabilized over time and non-rigid updates are limited by mechanical constraints. At the same time, the computational load remained low enough for real-time execution, addressing a common trade-off where stronger deformation modeling increases processing time. These results support a practical strategy for deformable AR navigation: improve depth consistency through temporal integration, then apply constrained non-rigid alignment to maintain anatomical realism while preserving real-time responsiveness.

4. Conclusion

This study introduced a temporal depth fusion–based markerless registration framework for deformable soft-tissue tracking in augmented reality surgical navigation. By combining multi-frame depth integration with physics-constrained non-rigid alignment, the method reduced depth noise and improved registration accuracy under controlled deformation conditions. The results showed a clear reduction in target registration error compared with rigid ICP and non-rigid ICP without mechanical constraints, while preserving real-time performance on a head-mounted device. These findings suggest that stabilizing depth measurements across consecutive frames and restricting deformation using biomechanical regularization can improve alignment stability in soft-tissue environments. The proposed framework addresses two common challenges in current AR navigation systems: sensitivity to noisy depth data and unrealistic surface deformation during matching. The observed millimeter-level accuracy indicates potential value for intraoperative guidance in procedures involving abdominal organs, where motion and manipulation are frequent. However, several limitations should be noted. Validation was performed on ex vivo specimens and phantom models rather than in vivo clinical cases. In addition, the mechanical model used for regularization was simplified and may not fully capture heterogeneous tissue properties. Further studies should include clinical evaluation, patient-specific parameter estimation, and system optimization to support practical deployment in surgical settings.

References

Wang, Y., Wen, Y., Wu, X., Wang, L., & Cai, H. (2024). Modulation of Gut Microbiota and Glucose Homeostasis through High-Fiber Dietary Intervention in Type 2 Diabetes Management.

- Asadi Ashenaabad, S., Lotfi, N., Daryabeigi Salami, D., Negareshifard, P., & Zarei, A. (2025). Bridging pre-operative planning and intra-operative decision-making: a systematic review of information-management strategies in AR-enhanced surgery. *InfoScience Trends*, 2(7), 12-25.
- 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*.
- Beachey, W. (2022). *Respiratory care anatomy and physiology E-book: Foundations for clinical practice*. Elsevier Health Sciences.
- Wang, Y., Wen, Y., Wu, X., & Cai, H. (2024). Comprehensive Evaluation of GLP1 Receptor Agonists in Modulating Inflammatory Pathways and Gut Microbiota.
- Gao, G., Gao, R., Gao, R., & Zhou, H. (2026). Engineering Analysis and Quantitative Research on the Platform-Based Evolution of Enterprise Communication Systems.
- Cvetić, J., Šekoranja, B., Švaco, M., & Šuligoj, F. (2025). Triplanar point cloud reconstruction of head skin surface from computed tomography images in markerless image-guided surgery. *Bioengineering*, 12(5), 498.
- Wang, Y., Yin, X., Chen, J., & Wang, Y. (2026). Evidence-Based Study on Low-Burden Digital Phenotyping for Precision Screening of Oral Anti-Obesity Drug Efficacy.
- Lucchetti, F., Céléreau, E., Steullet, P., Alemán-Gómez, Y., Hagmann, P., Klauser, A., & Klauser, P. (2025). Constructing the human brain metabolic connectome with MR spectroscopic imaging reveals cerebral biochemical organization. *Nature Communications*.
- Zhang, Y., Gu, W., & Wang, J. (2025). Research on First Article Inspection (FAI)-Driven Quality Assurance Methods for Wind Turbine Installation and Operation & Maintenance and Their Effect on Reliability Improvement. Available at SSRN 6094206.
- Maharjan, S., Inadomi, S., Itakura, K., & Chun, P. J. (2025). Domain-adaptive self-supervised learning for corrosion detection and 3D building information model mapping in steel tunnels. *Computer-Aided Civil and Infrastructure Engineering*, 40(26), 4425-4447.
- Zhang, Y., Gu, W., & Wang, J. (2026). Construction of Wind Farm Asset Health Index Based on Multi-Dimensional Indicators and Analytic Hierarchy Process and Its Correlation with Operational Performance. *Authorea Preprints*.
- Nivethika, S. D., Joshi, S., Verma, K., Aishwarya, V., Srinivasan, V. V., Pandian, M. S., & Paulraj, P. (2026). Attention-guided saptio-temporal feature fusion for robus video surveillance anomaly detection. *Scientific Reports*.
- Jiao, Y., Zhao, B., Wang, A., & Shi, T. (2026). Construction and Empirical Study of a Modularized Teaching System for Art Courses Based on a Unified Training Pathway.
- Assis, T. M. D. S. (2026). Physics-Guided Deep Learning for Sparse Data-Driven Brain Shift Registration.
- Jiao, Y., Wang, A., Zhao, B., & Shi, T. (2026). The Impact of Visual Language Strategies in Public Art Creation on Community Spatial Perception and Public Behavior.
- Oun, A., Hagerdorn, N., Scheideger, C., & Cheng, X. (2024). Mobile devices or head-mounted displays: a comparative review and analysis of augmented reality in healthcare. *IEEE Access*, 12, 21825-21839.
- Liu, S., Liu, X., & Feng, H. (2025, November). Research on AI-Driven Visual Design and Immersive Interactive Experiences Based on Multimodal Cognition and User. In *Proceedings of the 2025 International Conference on Digital Society and Intelligent Computing* (pp. 734-740).
- Ozturk, C., Pak, D. H., Rosalia, L., Goswami, D., Robakowski, M. E., McKay, R., ... & Roche, E. T. (2025). AI-Powered Multimodal Modeling of Personalized Hemodynamics in Aortic Stenosis. *Advanced Science*, 12(5), 2404755.
- Yu, X. (2026). Enhancing Search Efficiency through LLM-Based User Memory Systems for Query Matching and Intent Modeling.
- Cho, K. W., Youm, S., & Shin, K. S. (2026). Force Feedback-Based Adaptive Sensor Fusion Algorithm for Precision Enhancement in VR-Based Medical Training Systems. *IEEE Access*.

- Xiao, H., Zhu, W., Qiu, Z., Zeng, Z., Zhang, S., & Xiao, R. (2025, September). Anti-Drosophila Habituation Clustering for Enhanced Anomaly Detection in Data Streams. In 2025 10th International Conference on Computer and Information Processing Technology (ISCIPT) (pp. 636-644). IEEE.
- Frisk, H., Burström, G., Persson, O., El-Hajj, V. G., Coronado, L., Hager, S., ... & Elmi-Terander, A. (2024). Automatic image registration on intraoperative CBCT compared to Surface Matching registration on preoperative CT for spinal navigation: accuracy and workflow. *International Journal of Computer Assisted Radiology and Surgery*, 19(4), 665-675.
- Chen, J., Xiao, H., Zeng, Z., Du, X., & Xiao, R. (2026). Diffusion-step attention consistency for multivariate time series anomaly detection. *Knowledge-Based Systems*, 115716.
- Küttner, M., Zitz, V., Ramesh, S., Beigl, M., & Röddiger, T. (2026). EarResp-ANS: Audio-Based On-Device Respiration Rate Estimation on Earphones with Adaptive Noise Suppression. arXiv preprint arXiv:2602.03549.
- Wang, Y., Wang, L., Wen, Y., Wu, X., & Cai, H. (2025). Precision-Engineered Nanocarriers for Targeted Treatment of Liver Fibrosis and Vascular Disorders.
- Ge, M., Wang, S., Yang, Y., & Yao, J. (2026). AGSM-CPA: Reliability-Aware Robustness for Rotation-Invariant Point Cloud Learning. *Mathematics*, 14(2), 278.