This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Deformable CT-US Registration via Anatomy-Aware Implicit Neural Representations

Agnieszka Lach[1,2], Magdalena Wysocki[2], Feng Li[2], Mohammad Farid Azampour[2], Benjamin D. Killeen[2], Felix Ginzinger[1], Mathias Braun[1], Philipp Steininger[1], Heinz Deutschmann[1], and Nassir Navab[2]

[1] medPhoton GmbH, Salzburg, Austria
[2] Computer Aided Medical Procedures, Technical University of Munich, Munich, Germany
agnieszka.lach@tum.de

Abstract. Slice-to-volume registration between ultrasound (US) and preoperative computed tomography (CT) imaging would enhance many minimally invasive interventions, for example by locating soft tissue structures intra-operatively that are discernible in CT. While optical tracking enables initial rigid registration, contact from the probe induces soft tissue deformations that inhibit accurate alignment. In this work, we introduce a deformable CT-ultrasound registration framework that incorporates anatomical priors derived from CT to improve registration under deformation. Rigid registration is first established using a robot-assisted optical tracking system, after which a deformable transformation is estimated using a sinusoidal implicit neural representation (SIREN) optimized per frame. Tissue stiffness is approximated from CT-based HU values and used as spatially varying regularization, suppressing deformation in rigid structures such as bone while allowing more flexibility in soft tissue. Two additional constraints capture the physics of probe contact: a contact-zone displacement prior that drives the displacement field to compress tissue below the probe face, and a fan-geometry regularization term based on beam direction and convex transducer field of view. Model parameters are optimized with a normalized gradient field (NGF). The proposed approach improves alignment over rigid initialisation by 17% and outperforms classical deformable baselines while maintaining near-zero topological folding.

Keywords: Deformable · Minimally invasive · Computer-assisted

1 Introduction

Tracked intra-operative ultrasound (US) provides real-time, quantitative information about soft tissue dynamics, enabling high precision guidance for interventional procedures [12]. To overcome US-specific limitations such as acoustic shadowing and shallow penetration, there is growing interest in combining US with preoperative and intra-operative computed tomography (CT) imaging [8, 9, 11, 17], thereby accessing 3D anatomical context dynamically. However, even with accurate rigid co-registration, *e.g.*, via optical tracking, spatial

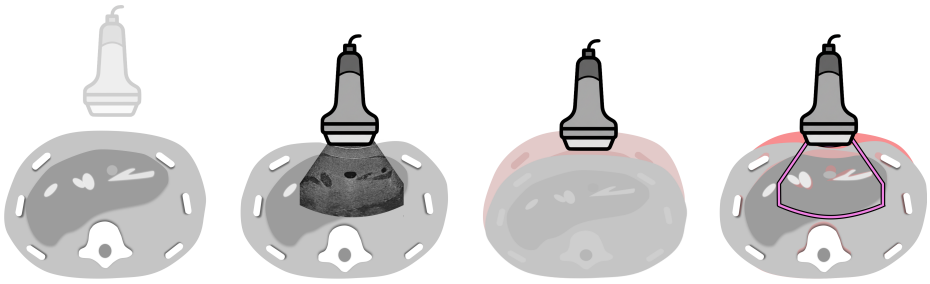

Fig. 1. To enable precise anatomical alignment between tracked ultrasound and CT (a), it is necessary to account for probe-induced soft-tissue deformation (b). Existing methods are failing to incorporate anatomical constraints, allowing for non-physically plausible deformations (c) (exaggerated for visual clarity). We propose a deformable registration framework that incorporates known physical constraints and anatomical priors to constrain the deformation field, improving alignment while maintaining physical plausibility (d).

misalignment persists due to probe-induced soft-tissue deformation during US acquisition. Strong contact between the probe and soft tissue is necessary to achieve sufficient image quality, but the resulting pressure introduces spatial inconsistencies between the US and CT volumes, as shown in Fig. 1a-b. This prevents precise anatomical alignment for high precision interventions such as liver tumor ablation or biopsy [17]. While electromagnetic tracking can estimate deformation directly [17], it introduces significant complexity to clinical workflows and may be susceptible to interference from metallic objects. Thus, there is a need for deformable registration methods that reconcile tracked US with CT volumes while accounting for probe-induced tissue deformation.

Existing methods lack anatomy-aware constraints and thus risk non-physically plausible deformations (Fig. 1c): physical priors are either impractical—e.g. Geng *et al.* [3] require contact-force sensors unavailable in CT-US—or, in recent learning-based work, only folded into the training objective of convolutional [1, 5] or implicit neural representation (INR) [13, 16] models. INRs offer a compelling per-pair alternative to amortized learning, motivating our anatomy-constrained INR formulation.

Here, we propose a CT-US registration framework that incorporates known physical constraints and anatomical priors into an INR of deformation. Our primary contribution is the method which uses three anatomy-aware loss terms tailored to the CT-US slice to volume probe-contact setting, which are integrated with a NGF-based multi-modal similarity metric and Jacobian regularization [16]. First, we encode tissue stiffness as a spatially varying regularization term derived from CT-based segmentation without requiring any finite-element solver. Second, we introduce a contact-zone displacement prior that enforces

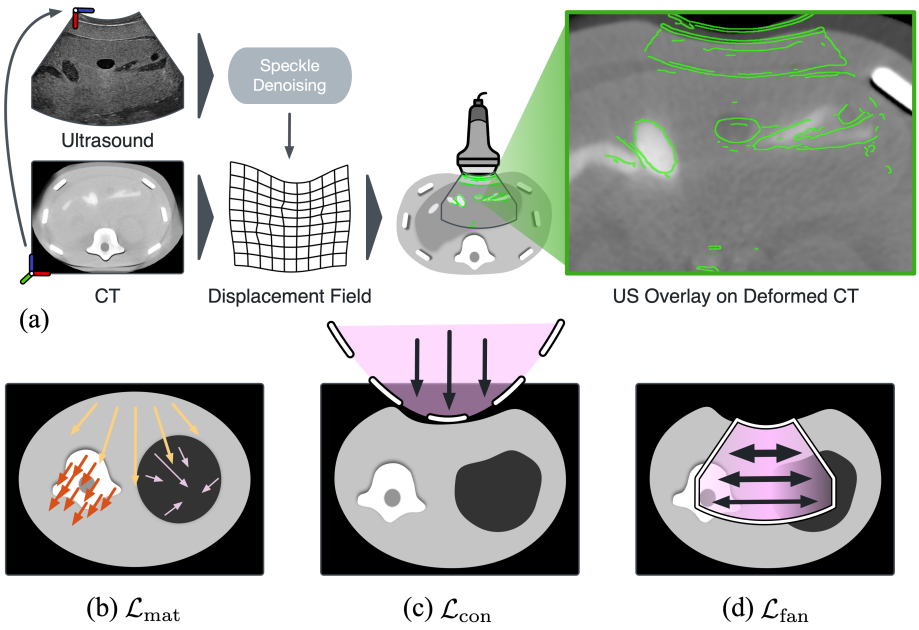

Fig. 2. Proposed non-rigid CT volume to US slice registration pipeline. The workflow transitions from initial rigid alignment to a displacement field estimation. The deformation is governed by three novel regularization terms: (a) $\mathcal{L}_{mat}$ for material stiffness preservation, (b) $\mathcal{L}_{con}$ for contact-zone interface modeling, and (c) $\mathcal{L}_{fan}$ for FOV-constrained spatial smoothing.

the physics of probe-induced surface compression by penalizing residual tissue in the crescent region between the probe face and the US fan arc. Finally, we tailor the optimization strategy to the fan-shaped, partial field of view of a convex transducer through beam-aligned directional regularization and an angular-tapered boundary weight that attenuates displacement toward the lateral fan horns where beam quality is weakest. We evaluate our approach on two datasets of robot-assisted US-CT acquisitions of abdominal phantoms, demonstrating superior alignment compared to both rigid initialization (up to 17% lower TRE) and classical deformable baselines.

2 Methods

Given a 2D US frame $I_{\text{US}} \in \mathbb{R}^{H \times W}$ acquired by a tracked convex transducer, we seek a dense displacement field $\mathbf{d}\colon \mathbb{R}^3 \to \mathbb{R}^3$ that maximizes the alignment of anatomical structures between the 3D CT volume I_{CT} and the US frame for all pixels $\mathbf{x}$ inside the US fan Ω_{fan}. Here, I_{CT} is resampled to the US pixel grid via optical-tracking calibration [8]. We parameterize $\mathbf{d}$ as a sinusoidal implicit neural representation (SIREN [14]) optimized *per frame*. Similarity between the

two modalities is measured with a normalized gradient fields (NGF) metric, and the Jacobian of $\mathbf{d}$ is regularized following Wolterink et al. [16]. As out-of-plane motion is unobservable from a single US slice, the displacement is constrained to the US imaging plane, yielding a 2.5-D warp.

2.1 CT-Derived Stiffness Regularization

We replace the spatially uniform regularization weights used in prior INR registration work with a per-voxel stiffness map derived semi-automatically from the CT volume. Each voxel receives a *continuous* stiffness scalar κ that follows its CT Hounsfield value and spans three orders of magnitude, from $\kappa \approx 0.01$ in air, through $\kappa \approx 1$ in soft tissue, to $\kappa \approx 100$ in bone, with morphological smoothing and hole-filling so that stiffness varies gradually across tissue interfaces rather than jumping at hard thresholds. This scalar κ_i acts as a continuous material mask that gates how rigidly each point moves with its neighborhood. Rather than penalizing displacement magnitude, we penalize each point's deviation from the κ-weighted *rigid* consensus motion $\mathbf{T}^\star \in SE(3)$ of the stiff structure, in proportion to κ_i:

$$\mathcal{L}_{\text{mat}} = \frac{\sum_i \kappa_i \big\|\mathbf{d}(\mathbf{x}_i) - \mathbf{T}^\star(\mathbf{x}_i)\big\|^2}{\sum_i \kappa_i}, \qquad \mathbf{T}^\star = \underset{\mathbf{T}\in SE(3)}{\arg\min} \sum_i \kappa_i \big\|\mathbf{d}(\mathbf{x}_i) - \mathbf{T}(\mathbf{x}_i)\big\|^2 \tag{1}$$

where $\mathbf{T}^\star$ is the κ-weighted least-squares rigid fit to the current field, solved in closed form each iteration. A bone point ($\kappa = 100$) is thus driven onto the rigid consensus—zero shape change, while remaining free to translate and rotate—whereas soft tissue ($\kappa \approx 1$) is left essentially unconstrained and air ($\kappa \approx 0$) drops out. A Jacobian regularizer $\mathcal{L}_{\text{reg}} = \frac{1}{N}\sum_i \big[\lambda_F \|\mathbf{J}_i\|_F^2 + \lambda_d\, w(\kappa_i)\,(\det(\mathbf{I} + \mathbf{J}_i) - 1)^2\big]$ is retained for global smoothness and fold prevention, where the folding (determinant) penalty is itself stiffness-weighted by $w(\kappa_i) = 1/(100.1 - \kappa_i)$ so that volume change is suppressed most strongly inside stiff structures.

2.2 Contact-Zone Displacement Prior

We introduce a novel loss term that encodes the physical reality of probe contact: when the probe is pressed against tissue, the crescent-shaped region between its flat face and the concave fan arc must be air. The extended domain Γ covers every pixel column between the upper fan horns, from the image top down to the fan's concave arc.

Displacement prior loss. The INR is queried at crescent coordinates $\mathbf{x} \in \Gamma$ each iteration, and the CT is sampled at the displaced positions:

$$\mathcal{L}_{\text{con}} = \frac{1}{|\Gamma|} \sum_{\mathbf{x}\in\Gamma} \Big[\text{ReLU}\big(I_{\text{CT}}(\mathbf{x} + \mathbf{d}(\mathbf{x})) - a_0\big)/s\Big]^2, \tag{2}$$

where a_0 and s are the air-intensity reference and scale of the normalized CT; the one-sided rectifier penalizes only above-air (tissue-level) intensity remaining in the crescent. Minimizing $\mathcal{L}_{\text{con}}$ drives the field to push CT coordinates toward low-intensity regions, enforcing downward tissue compression without any ground-truth annotation.

2.3 Fan-Geometry-Aware Optimization

We tailor the optimization strategy to the fan-shaped, partial field of view of a convex transducer through three geometric mechanisms: fan-constrained patch sampling, beam-aligned directional regularization, and a smooth boundary weight.

Fan-constrained patch sampling. Patch centers are drawn uniformly from within Ω_{fan}, ensuring every queried location has valid spatial neighbors for the NGF gradient computation and no network capacity is spent on background.

Beam-aligned directional regularization. For each pixel inside the fan, a unit beam vector $\hat{\mathbf{b}}(\mathbf{x})$ is computed from the predefined fan apex to $\mathbf{x}$, and the in-plane displacement is decomposed into radial (along-beam) and tangential (cross-beam) parts. An asymmetric penalty

$$\mathcal{L}_{\text{fan}} = \frac{\lambda_{\text{tan}}}{N} \sum_i \left\| \mathbf{d}^i_{\text{tan}} \right\|^2 + \frac{1}{N} \sum_i \Big(\lambda_r \, \text{ReLU}(d^i_{\text{rad}})^2 + \lambda_r^- \, \text{ReLU}(-d^i_{\text{rad}} - \tau)^2 \Big) \quad (3)$$

encodes the physical prior that the probe acts as a compressor: retrograde motion (toward the apex) is penalized more strongly than forward compression, which is allowed freely; a small deadzone τ leaves minor toward-probe motion unpenalized.

Smooth boundary weight. Predicted displacements are multiplied element-wise by a precomputed map $m(\mathbf{x}) \in [0, 1]$, constructed by dilating the fan mask outward, applying a Gaussian ramp, and multiplying by an angular taper $\cos^2(\theta/\theta_{\max})$ that attenuates displacement toward the lateral fan horns where beam quality is weakest.

2.4 Implementation Details

To enhance structural features for the similarity metric, both modalities undergo intensity normalization, and the ultrasound frames are preprocessed using the self-supervised Speckle2Self [10] denoising framework to mitigate signal noise and acoustic artifacts. The SIREN backbone uses 2 layers and 128 hidden units. Raw voxel indices are min-max normalized to $[0, 1]^3$ before network input; outputs are rescaled to voxel units before warping via trilinear interpolation [14, 16]. The full objective is $\mathcal{L} = \lambda_{\text{NGF}} \mathcal{L}_{\text{NGF}} + \mathcal{L}_{\text{mat}} + \mathcal{L}_{\text{fan}} + \mathcal{L}_{\text{con}} + \mathcal{L}_{\text{reg}} + \mathcal{L}_{\text{band}}$, where $\mathcal{L}_{\text{reg}}$ is the Jacobian regularizer above and $\mathcal{L}_{\text{band}}$ applies the same smoothness

penalty to a dilated band just outside the fan, keeping the out-of-view extrapolation continuous and preventing off-plane rigid structures from deforming. It is optimized with Adam (lr 3×10^{-5}, gradient clip 1.0) for up to 200 iterations. Validation NGF is evaluated every 50 iterations; optimization stops early if no improvement exceeds 10^{-4} for 50 checks, restoring best-seen weights.

3 Experiments and Results

We evaluate on four robotic US–CT sweeps spanning two acquisition foci: **Dataset 1**, acquired on a CIRS Model 057A Triple Modality Abdominal Phantom, and **Dataset 2**, acquired on a Kyoto Kagaku Dual Modality Human Abdomen Phantom (Model US-22); both were acquired with a robot-assisted setup [8]. Manual landmarks were annotated on five evenly spaced frames per sweep (478 pairs total), and the deformation is run per-frame on the full cine. Ultrasound (US) data were captured at 880×660 resolution using a Siemens ACUSON Juniper (5C1 probe). For CT, a mobile "ImagingRing" (medPhoton, Austria) was used. The robotic systems were co-registered via hand-eye calibration using the Ring's integrated optical tracker.

Metrics. Registration accuracy is measured with the *(i)* Target Registration Error (TRE), the mean Euclidean distance between corresponding manually placed landmarks after registration (478 landmark pairs across all sweeps; significance vs. rigid by Wilcoxon signed-rank test). Image agreement is measured by *(ii)* phase-congruency NCC (PC) [7], a contrast-invariant US–CT similarity measure (higher = better); we deliberately use a phase-based rather than a gradient-based similarity metric here, as the latter mirrors our NGF data term and would bias the evaluation in our favor. Plausibility is measured by *(iii)* the Jacobian folding fraction, the percentage of pixels with $\det(\mathbf{I} + \mathbf{J}) < 0$ (lower = better, zero is ideal), and—on the bone-adjacent sweeps—*(iv)* the bone/soft strain ratio, the mean shape-change $\|\nabla \mathbf{d}\|$ inside bone relative to soft tissue (lower = bone held more rigid; 1 = bone deforms like soft). As a secondary appearance measure we additionally report the Dice overlap of a set of corresponding anatomical structures delineated in both modalities.

Baselines. We compare against *(i)* rigid initialization only (no deformable stage); *(ii)* Demons [15], a classical diffeomorphic method; *(iii)* B-spline FFD [6] with two data terms, local cross-correlation (FFD-LCC) and normalized gradient fields (FFD-NGF); and *(iv)* Farnebäck optical flow. To justify our NGF data term, we also report our own network driven instead by *(v)* MIND-only [4] and *(vi)* LC2-only [2]. All methods share identical rigid pre-alignment and run per-frame.

3.1 Quantitative Results

Tab. 1 summarizes performance pooled across all four sweeps (478 landmarks). Our method attains the lowest TRE (3.97 mm; $p < 10^{-10}$ vs. rigid), outperforming the strongest baseline (FFD-NGF, 4.43 mm) and all others, while producing

Table 1. Quantitative comparison, pooled over all four sweeps (478 landmarks). The p column reports a Wilcoxon signed-rank test vs. rigid; all deformable methods improve significantly ($p < 0.05$) except MIND-/LC2-only (n.s.), while Farnebäck is significantly *worse*. PC: phase-congruency NCC. Best in each column in **bold**, second best underlined.

Method	PC ↑	Dice ↑	Fold% ↓	p vs. rigid	TRE [mm] ↓
Rigid only	–	–	—	–	4.79
Demons [15]	0.045	0.56	2.39	1×10^{-7}	4.59
FFD-LCC [6]	0.032	0.54	0.29	0.02	4.50
FFD-NGF [6]	**0.053**	**0.65**	1.41	7×10^{-5}	4.43
Farnebäck	0.012	0.45	5.58	4×10^{-22}	5.85
MIND-only [4]	0.037	0.48	0.03	0.08	4.63
LC2-only [2]	0.025	0.48	**0.02**	0.88	4.80
Ours	0.048	0.55	0.03	6×10^{-14}	**3.97**

essentially fold-free deformations (0.03% vs. 1.4–5.6% for FFD/Demons/optical-flow). FFD-NGF attains a marginally higher phase-congruency similarity (0.053 vs. our 0.048) and a higher Dice by deforming more aggressively—but at the cost of pervasive folding (1.41%) and with no notion of tissue material, i.e. greater appearance similarity without better landmark accuracy. Replacing our NGF data term with MIND or LC2 loses all significance against rigid ($p = 0.08$ and 0.88), confirming NGF as the appropriate US–CT driver. Fig. 3 shows qualitative results on soft-tissue and bone-adjacent frames.

3.2 Ablation Study

Tab. 2 isolates each component. Removing the material term $\mathcal{L}_{\text{mat}}$ leaves TRE essentially unchanged (3.95 vs. 3.97 mm) yet the bone/soft strain ratio jumps from 0.58 to 1.10: the material term enforces bone rigidity *at no accuracy cost*, which is precisely its intended role—soft tissue deforms while the rib is held rigid. Removing the contact/displacement prior $\mathcal{L}_{\text{con}}$ degrades TRE by 0.21 mm and lowers phase-congruency similarity. $\mathcal{L}_{\text{fan}}$ is TRE-neutral, as it constrains only the near-field region that contains no landmarks. Finally, substituting the NGF data term with MIND or LC2 collapses accuracy to no-better-than-rigid (4.63/4.80 mm, n.s.), the strongest evidence for the NGF choice. Landmarks were placed by a clinically trained annotator (∼20–29 points per frame).

4 Conclusion

We presented a deformable CT-US registration framework based on implicit neural representations of the displacement field, augmented with three anatomy-aware constraints. Together these components address the two principal failure modes of existing INR registration methods in the CT–US setting, which are non-physical deformation of rigid structures and unconstrained behavior outside the imaging field of view. Our results demonstrate improved alignment over

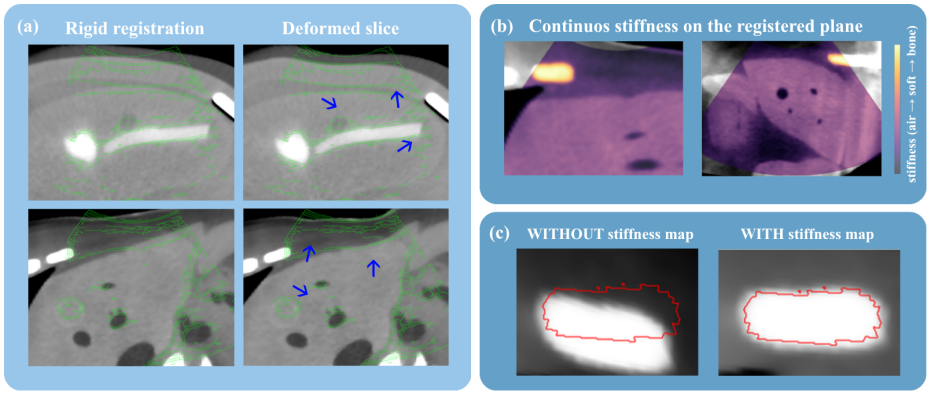

Fig. 3. (a) Qualitative assessment of the stiffness-constrained deformable registration. (b) By incorporating a stiffness map context, the model corrects for transducer-induced deformation, providing the sub-millimeter structural alignment necessary for guided interventions and planning. (c) Note the stability of the bone relative to the compliant soft tissue.

rigid initialization and classical deformable baselines on both soft-tissue and bone-containing phantoms, with near-zero topological folding throughout. Several limitations remain. The stiffness map relies on a CT-intensity-to-stiffness mapping calibrated for the phantom that may generalize less well to pathological tissue or implants, and the fan-geometry assumptions are specific to convex transducers. A significant challenge in US-CT registration is the limited reliability of similarity metrics, which often fail to capture complex morphological correspondences. To address this, we relied on expert-validated anatomical landmarks to calculate the TRE. While our method demonstrates high qualitative and quantitative alignment with phantom structures, we acknowledge that the transition to *in vivo* studies may introduce greater anatomical variability. A ro-

Table 2. Ablation of the loss terms and the data term (pooled over all sweeps; bone/soft strain on the bone-adjacent sweeps). PC: phase-congruency NCC; p: Wilcoxon vs. rigid. The material term $\mathcal{L}_{\text{mat}}$ drives the bone/soft ratio well below 1 at no TRE cost; removing it returns the ratio to ≈ 1 (bone deforms like soft). Best in each column in **bold**, second best underlined.

Configuration	PC ↑	bone/soft ↓	Fold% ↓	p vs. rigid	TRE [mm] ↓
w/o $\mathcal{L}_{\text{mat}}$	**0.049**	1.10	0.03	6×10^{-15}	**3.95**
w/o $\mathcal{L}_{\text{con}}$	0.036	**0.43**	**0.00**	7×10^{-13}	4.18
w/o $\mathcal{L}_{\text{fan}}$	0.048	0.60	0.03	5×10^{-13}	3.96
MIND data term	0.037	–	0.03	0.08	4.63
LC2 data term	0.025	–	0.02	0.88	4.80
Full model (ours)	0.048	0.58	0.03	6×10^{-14}	3.97

bust, modality-independent registration metric is needed to reduce reliance on manual landmark selection and mitigate potential inter-observer bias, and while probe placement makes simultaneous CT-US capture challenging, future work would benefit from ground truth volumetric deformations based on volumetric imaging for comprehensive validation.

Bibliography

[1] Azampour, M., Tirindelli, M., Lameski, J., Gafencu, M., Tagliabue, E., Fatemizadeh, E., Hacihaliloglu, I., Navab, N.: Anatomy-aware computed tomography-to-ultrasound spine registration. Medical Physics **51** (09 2023). https://doi.org/10.1002/mp.16731

[2] Fuerst, B., Wein, W., Müller, M., Navab, N.: Automatic ultrasound–mri registration for neurosurgery using the 2d and 3d lc2 metric. Medical Image Analysis **18**(8), 1312–1319 (2014). https://doi.org/https://doi.org/10.1016/j.media.2014.04.008, https://www.sciencedirect.com/science/article/pii/S1361841514000620, special Issue on the 2013 Conference on Medical Image Computing and Computer Assisted Intervention

[3] Geng, Y., Zhao, M., Xu, F., Cao, G., Meng, G., Liu, H.: PADReg: Physics-aware deformable registration guided by contact force for ultrasound sequences. arXiv preprint arXiv:2508.08685 (2025). https://doi.org/10.48550/arXiv.2508.08685

[4] Heinrich, M.P., Jenkinson, M., Bhushan, M., Matin, T., Gleeson, F.V., Brady, S.M., Schnabel, J.A.: Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. Medical Image Analysis **16**(7), 1423–1435 (2012). https://doi.org/https://doi.org/10.1016/j.media.2012.05.008, https://www.sciencedirect.com/science/article/pii/S1361841512000643, special Issue on the 2011 Conference on Medical Image Computing and Computer Assisted Intervention

[5] Jian, B., Azampour, M.F., De Benetti, F., Oberreuter, J., Bukas, C., Gersing, A.S., Foreman, S.C., Dietrich, A.S., Rischewski, J., Kirschke, J.S., Navab, N., Wendler, T.: Weakly-supervised biomechanically-constrained ct/mri registration of the spine. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VI. p. 227–236. Springer-Verlag, Berlin, Heidelberg (2022). https://doi.org/10.1007/978-3-031-16446-0_22, https://doi.org/10.1007/978-3-031-16446-0_22

[6] Klein, S., Staring, M., Murphy, K., Viergever, M., Pluim, J.: Elastix: A toolbox for intensity-based medical image registration. IEEE transactions on medical imaging **29**, 196–205 (11 2009). https://doi.org/10.1109/TMI.2009.2035616

[7] Kovesi, P.: Image features from phase congruency. Videre: Journal of Computer Vision Research **1**(3), 1–26 (1999)

[8] Li, F., Bi, Y., Huang, D., Jiang, Z., Navab, N.: Robotic CBCT meets robotic ultrasound. International Journal of Computer Assisted Radiology and Surgery **20**, 1049–1057 (2025). https://doi.org/10.1007/s11548-025-03336-x

[9] Li, F., Li, Z., Jiang, Z., Navab, N., Bi, Y.: Robotic ultrasound makes CBCT alive. arXiv preprint arXiv:2603.10220 (2026)

[10] Li, X., Navab, N., Jiang, Z.: Speckle2self: Self-supervised ultrasound speckle reduction without clean data. Medical Image Analysis **106**, 103755 (2025). https://doi.org/https://doi.org/10.1016/j.media.2025.103755, https://www.sciencedirect.com/science/article/pii/S1361841525003020

[11] Masoumi, N., Belasso, C.J., Ahmad, M.O., Benali, H., Xiao, Y., Rivaz, H.: Multimodal 3D ultrasound and CT in image-guided spinal surgery: public database and new registration algorithms. Int. J. CARS **16**(4), 555–565 (Apr 2021). https://doi.org/10.1007/s11548-021-02323-2

[12] Pavone, M., Seeliger, B., Teodorico, E., Goglia, M., Taliento, C., Bizzarri, N., Lecointre, L., Akladios, C., Forgione, A., Scambia, G., Marescaux, J., Testa, A.C., Querleu, D.: Ultrasound-guided robotic surgical procedures: a systematic review. Surg. Endosc. **38**(5), 2359–2370 (May 2024). https://doi.org/10.1007/s00464-024-10772-4

[13] Sideri-Lampretsa, V., McGinnis, J., Qiu, H., Paschali, M., Simson, W., Rueckert, D.: Sinr: Spline-enhanced implicit neural representation for multimodal registration. Proceedings of Machine Learning Research (2024)

[14] Sitzmann, V., Martel, J.N.P., Bergman, A.W., Lindell, D.B., Wetzstein, G.: Implicit neural representations with periodic activation functions. CoRR **abs/2006.09661** (2020), https://arxiv.org/abs/2006.09661

[15] Thirion, J.P.: Image matching as a diffusion process: an analogy with maxwell's demons. Medical Image Analysis **2**(3), 243–260 (1998). https://doi.org/https://doi.org/10.1016/S1361-8415(98)80022-4, https://www.sciencedirect.com/science/article/pii/S1361841598800224

[16] Wolterink, J.M., Zwienenberg, J.C., Brune, C.: Implicit neural representations for deformable image registration. In: Proceedings of the International Conference on Medical Imaging with Deep Learning (MIDL). pp. 1349–1359 (2022)

[17] Xu, Z.F., Xie, X.Y., Kuang, M., Liu, G.J., Chen, L.D., Zheng, Y.L., Lu, M.D.: Percutaneous radiofrequency ablation of malignant liver tumors with ultrasound and CT fusion imaging guidance. J. Clin. Ultrasound **42**(6), 321–330 (Jul 2014). https://doi.org/10.1002/jcu.22141