SRP99 Final Report: 3D Point Cloud Intra-Patient Lung Registration via Optimal Transport Methods

Chou (Athena) Mo[a]

[a]Department of Mathematics, University of California - Los Angeles, Los Angeles, CA, USA

ABSTRACT

This report is for the course Radiological Science 99 (instructed by Dr. William Hsu and guided by Yunzheng Zhu). Accurate intra-patient lung CT registration is critical for tracking disease progression, evaluating treatment response, and other biomedical applications. However, it remains challenging due to complex non-rigid motion and large deformations from respiration. Our work this quarter focuses on recreating the experiments detailed in the study titled "Iterative optimal transport for multimodal image registration" for 3D point cloud volumes.[1]

1. INTRODUCTION

Accurate intra-patient registration of lung images provides key information for evaluating patients' respiratory motion, marking anatomical landmarks, and planning in thoracic radiotherapy.[2–4]

In previous studies, dense CT volumes have been investigated using intensity-based similarity measures (combined with spline-field or velocity-field parameterizations) for 3D deformable registration. However, performance robustness can be diminished due to low soft-tissue contrast, complex sliding motion, and large differences between inspiratory and expiratory scans.[5–7] Motivated by these challenges, studies have been conducted on point-cloud based registration for derived anatomical structures such as vessels and airways, which is a better representation of anatomical topology.[2,8,9]

Optimal transport (OT) provides a mathematical framework for measuring distances between distributions and computing correspondences, and has therefore become a popular tool for both image and point-cloud registration.[1,6,10] This project proposes a 3D intra-patient lung registration method that adopts the iterative optimal transport (IOT) framework from using 2D multimodal edge images to 3D pulmonary vessel point clouds derived from paired inspiratory and expiratory CT scans. By iteratively updating the transport plan and transformation, IOT is designed to be robust to incomplete vessel extraction, anatomical differences across respiratory phases, and outliers arising from segmentation noise, with the ultimate goal of improving downstream tasks such as motion-compensated dose computation and accurate longitudinal tracking of lung lesions.[1,2,8]

1.1 Related Works

Existing multimodal image registration methods can be categorized as intensity-based or feature-based approaches.[4,11] Under a chosen deformation model, intensity-based methods optimize similarity measures, but they often degrade under severe deformations, partial overlap, or strong appearance differences, which are all typical for inspiratory–expiratory lung CT.[3,4] On the other hand, feature-based methods align salient structures using descriptors (e.g., RGF, DASC, RIFT) to improve robustness for modality changes, but are largely dependent on reliable feature extraction and sufficient anatomical overlap.[12,13]

Beyond classical strategies, deep learning–based registration methods learn both feature representations and deformation models. Previous work has shown strong performance on several benchmarks, but it requires large collections of aligned training pairs or landmarks, which are often lacking in clinical lung imaging.[3,4,14] For intra-patient lung registration, datasets such as Lung250M-4B provide paired inspiratory–expiratory CT scans with associated vessel point clouds. This enables systematic evaluation of both image-based and point-cloud based methods and motivates registration algorithms that can cope with strong deformations and differences in vessel visibility.[2,8]

Optimal transport offers a complementary geometric constraint for registration. OT has been used in mass-preserving formulations for dense image registration, and subsequent work has developed methods such as Unbalanced OT (UOT) that allow partial matching when structures appear or disappear between scans.[7,10,15] In

vascular and point-cloud scenarios, OT-based methods have shown improved robustness to intensity variability, outliers, and partial overlap, and can be integrated into a variety of pipelines.[8,9] The iterative optimal transport framework of Li et al. combines unbalanced OT with a polynomial transformation, achieving strong performance on challenging multimodal medical image datasets. Our work adapts this idea to 3D pulmonary vessel point clouds for intra-patient lung registration.[1,15]

2. METHOD

2.1 Overview

We build on the Iterative Optimal Transport (IOT) framework proposed by Li et al.,[1] which formulates image registration as a sequence of unbalanced optimal transport (UOT) problems on a set of structural points. The core algorithm alternates between two steps until convergence: (i) estimating a correspondence matrix $\mathbf{T}$ via UOT, and (ii) refining a polynomial transformation f via quasi-Newton optimization. We adapt this framework from 2D multimodal image registration to the problem of 3D lung CT point cloud registration between expiratory (EXP) and inspiratory (INSP) phases of the PVT1010 dataset.[8]

2.2 IOT

Given a moving point set $\{x_i\}_{i=1}^n \subset \mathbb{R}^3$ and a fixed point set $\{y_j\}_{j=1}^m \subset \mathbb{R}^3$, IOT minimizes the regularized UOT objective:

$$\min_{f \in \mathcal{H}} \mathrm{UOT}_\lambda(f) + \varepsilon R(f), \tag{1}$$

$$\mathrm{UOT}_\lambda(f) := \min_{\mathbf{T} \in \mathbb{R}_+^{n \times m}} \sum_{i,j} \|f(x_i) - y_j\|^2 T_{ij} + \lambda\, \mathrm{KL}(\mathbf{T}\mathbf{1}_m \| \mathbf{a}) + \lambda\, \mathrm{KL}(\mathbf{T}^\top \mathbf{1}_n \| \mathbf{b}), \tag{2}$$

where $\mathbf{T} \in \mathbb{R}_+^{n \times m}$ is the transport plan with T_{ij} denoting the mass transported from x_i to y_j; $\mathbf{1}_m \in \mathbb{R}^m$ and $\mathbf{1}_n \in \mathbb{R}^n$ are all-ones vectors; $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$ are the source and target marginal distributions; and $\mathrm{KL}(\mathbf{p}\|\mathbf{q}) = \sum_i p_i \log(p_i/q_i) - p_i + q_i$ is the Kullback–Leibler divergence with $R(f) = \sum_i \|f(x_i) - x_i\|^2$ penalizing excessive displacement from the initial positions. Transformation f is modeled as a global polynomial of degree q. The transport plan $\mathbf{T}$ is updated using the maximization-minimization solver of Chapel et al.,[16] and f is updated via BFGS. We use $q = 2$, $\varepsilon = 10^{-3}$, and search over $\lambda \in \{10^{-3}, 5 \times 10^{-3}, 2 \times 10^{-2}, 10^{-1}, 5 \times 10^{-1}\}$, selecting the value that minimizes mean nearest-neighbor distance on the validation set.

2.3 Adaptations

2.3.1 Point sampling

Instead of the 2D edge-point sampler that is used in the original IOT paper, we introduce a 3D structure-tensor corner detector adapted from the classical Förstner operator.[17] For each point p_i, we construct the 3×3 structure tensor over its k nearest neighbours:

$$S_i = \sum_{j \in \mathcal{N}(i)} (p_j - p_i)(p_j - p_i)^\top, \tag{3}$$

and compute the cornerness score:

$$w_i = \frac{\det(S_i)}{\mathrm{tr}(S_i)^3 + \epsilon}. \tag{4}$$

Points with high w_i correspond to locally isotropic neighborhoods. In the `foerstner_scalar` variant, w_i is further weighted by the normalized `radius` scalar to bias sampling toward vascular or peripheral structures as required. We keep the top $N = 1000$ points from each phase.

2.3.2 Point Normalization

The original IOT normalizes point sets by a single global scale factor. Since we adopt 3D point clouds, we instead apply per-axis normalization. We subtract the per-axis mean and divide by the per-axis standard deviation independently for x, y, and z.

2.3.3 Removal of KL

We also evaluate a balanced OT variant (`USE_KL=False`), where the KL marginal penalty terms (see Equation (2)) are removed entirely by setting λ to zero, and the transport plan is computed via the exact Earth Mover's Distance (EMD; `ot.emd`), which finds the minimum-cost coupling under strict mass conservation constraints.[18] This is to assess whether the unbalanced relaxation — adopted in the original paper due to considerations for modality-induced partial overlaps in 2D images — is still necessary in the same-modality 3D setting.

2.3.4 Evaluation

The original IOT paper evaluates performance by comparing against manually labeled landmark pairs using RMSE. However, the PVT1010 dataset that we use does not provide landmark annotations. Therefore, we evaluate registration performance using mean, median, and maximum nearest-neighbor distance from the warped EXP point cloud to the INSP point cloud, computed in millimeters after de-normalization.

3. EXPERIMENTS

3.1 Data and Experimental Setup

Experiments were conducted on the PVT1010 lung CT dataset using case `copd_000001`, a subject with chronic obstructive pulmonary disease (COPD).[19] Each scan phase is represented as a 3D point cloud extracted from the vascular tree and airway structure of the lung, stored as a VTK mesh with two scalar fields per point: a Hessian-based vesselness score and an estimated local tube radius in millimeters. Registration is performed between the expiratory (EXP, *moving*) and inspiratory (INSP, *fixed*) phases.

To investigate the effect of different point-cloud sampling strategies, polynomial degree, and optimal-transport formulation, we evaluate eight configurations from the combination of three factors:

- **Sampler** – farthest-point sampling (FPS) vs. Förstner-scalar sampling (FS)
- **Polynomial degree** – $d \in \{2, 3\}$ for the displacement basis $\Phi(X)$.
- **Transport formulation** – unbalanced OT with KL marginal penalties (UOT-KL, the original IOT) vs. balanced exact EMD (no-KL).

For each configuration, N=1000 points are sampled from each phase. The regularization weight is fixed at ϵ=10^{-3} and the outer alternating-minimization loop runs for at most 100 iterations with convergence threshold 10^{-6}. For UOT-KL runs, the marginal-penalty parameter λ is selected by grid search over $\{10^{-3}, 5{\times}10^{-3}, 2{\times}10^{-2}, 10^{-1}, 5{\times}10^{-1}\}$, and the value yielding the lowest post-registration error is retained.

3.2 Evaluation and Testing Method

Registration quality is measured by the mean nearest-neighbour (NN) distance from the warped EXP cloud to the INSP cloud:

$$\mathcal{E} = \frac{1}{n} \sum_{i=1}^{n} \min_j \|f(x_i) - y_j\|_2, \tag{5}$$

where $f(x_i) = \Phi(x_i)\, T$ is the polynomial warp applied to the i-th EXP point and $\{y_j\}$ are the INSP points. We report $\mathcal{E}$ both before registration (raw EXP-to-INSP distance, serving as the baseline) and after registration, together with the percentage improvement:

$$\Delta\mathcal{E}(\%) = \left(1 - \frac{\mathcal{E}_{\text{after}}}{\mathcal{E}_{\text{before}}}\right) \times 100. \tag{6}$$

In addition, we record the mean and maximum per-point displacement $\|f(x_i) - x_i\|_2$ to quantify the magnitude of the learned deformation field.

4. RESULTS

Quantitative results for all eight configurations are summarized in Table 1 and visualized in Figure 1.

Table 1. Grid search results for IOT registration on `copd_000001`. Post-registration mean NN distance reported as mean ± std (mm). Balanced-OT runs (noKL) have no λ parameter (N/A).

Sampler	**Deg**	**KL**	λ	**Pre-mean**	**Post-mean ± std**	**Improv.%**	**Disp-mean**	**Time (s)**
Farthest-Point Sampling (FPS)								
FPS	2	✓	0.02	9.382	8.823 ± 2.973	6.0%	21.047	633.4
FPS	2	×	N/A	9.382	9.175 ± 3.082	2.2%	25.411	2.6
FPS	3	✓	0.02	9.382	8.729 ± 3.128	7.0%	23.383	1064.4
FPS	3	×	N/A	9.382	9.140 ± 2.966	2.6%	25.314	2.5
Förstner (radius-weighted)								
Förstner	2	✓	0.005	10.812	4.402 ± 3.050	59.3%	29.071	535.1
Förstner	2	×	N/A	10.812	7.733 ± 4.759	28.5%	30.994	3.6
Förstner	3	✓	0.02	10.812	**4.335 ± 3.099**	**59.9%**	28.951	736.6
Förstner	3	×	N/A	10.812	7.391 ± 3.710	31.6%	30.330	2.8

Table 2. Paired t-tests on per-point post-registration NN distances across three comparison families. Each pair is matched on all other axes (e.g., sampler comparisons are matched on degree and KL setting).

Comparison	t**-statistic**	p**-value**	**Sig.**
(1) Sampler: FPS vs. Förstner (radius-weighted)			
deg= 2, KL	32.356	1.11×10^{-157}	***
deg= 2, noKL	7.975	4.15×10^{-15}	***
deg= 3, KL	30.141	1.81×10^{-142}	***
deg= 3, noKL	11.539	5.29×10^{-29}	***
(2) Transport: UOT-KL vs. Balanced-OT (noKL)			
FPS, deg= 2	−3.277	1.09×10^{-3}	**
FPS, deg= 3	−3.701	2.27×10^{-4}	***
Förstner, deg= 2	−20.950	4.52×10^{-81}	***
Förstner, deg= 3	−25.086	4.28×10^{-108}	***
(3) Polynomial degree: deg= 2 vs. deg= 3			
FPS, KL	0.934	3.50×10^{-1}	ns
FPS, noKL	0.444	6.57×10^{-1}	ns
Förstner, KL	1.261	2.08×10^{-1}	ns
Förstner, noKL	3.628	3.00×10^{-4}	***

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; ns = not significant. Paired t-test on per-point NN distances ($n = 1000$).

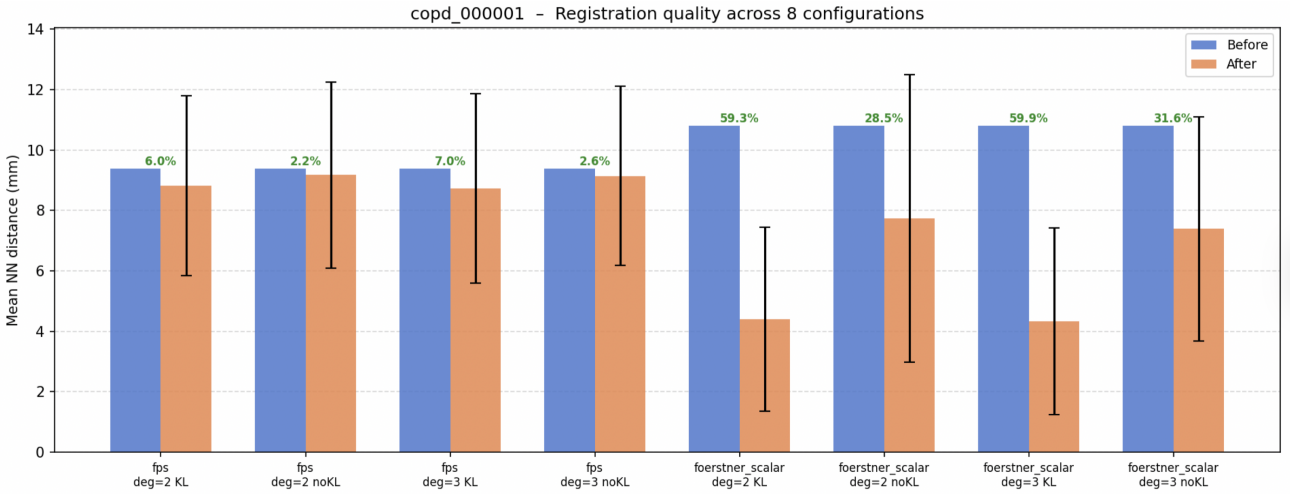

Figure 1. Mean nearest-neighbour distance before (blue) and after (orange) registration across all eight configurations on `copd_000001`. Percentage improvement over baseline is annotated above each after-registration bar.

The Förstner-scalar sampler outperforms FPS across all degrees and OT combinations. Despite starting from a higher baseline misalignment (≈10.8 mm vs. ≈9.4 mm), Förstner-scalar reduces the mean NN distance to ≈4.3–7.7 mm, compared with a best of 8.7 mm for FPS, suggesting that radius-weighted corner detection selects points with more effective local geometry. Within the Förstner-scalar group, UOT-KL outperforms balanced IOT by a wide margin (≈59% vs. ≈28–32% improvement), indicating that the KL marginal penalties play a critical role in handling the mass imbalance between fixed and moving lung point clouds.

Increasing the polynomial degree from 2 to 3 yields only marginal additional improvement under UOT-KL (FPS: 6.0% → 7.0%; Förstner: 59.3% → 59.9%) and a negligible effect for balanced IOT, suggesting that the higher-order basis provides little benefit beyond what is already achieved by the KL penalties. The best overall configuration is Förstner-scalar sampling, degree-3 polynomial basis, and UOT-KL, which achieves a post-registration mean NN distance of 4.335 mm, a 59.90% reduction from the unregistered baseline.

5. DISCUSSION

The dominant factor across all configurations for better registration is the point sampler. The Förstner cornerness score selects geometrically informative points that provide more stable support for the transport plan than spatially uniform but geometrically agnostic FPS points. Specifically, by design, Förstner selects locally isotropic regions where vessel branches bifurcate.

Although the original IOT paper primarily employs UOT-KL penalties via modality-induced partial overlap in 2D images, the KL marginal terms are equally important in the same modality 3D setting. Paired EXP and INSP scans show substantial mass imbalance, in which vessels visible in one respiratory phase may collapse or be absent in the other due to phase-dependent vascular visibility and segmentation noise, making strict mass conservation rigid.

The large standard deviation (∼3 mm) reflects spatially non-uniform registration quality. The polynomial warp captures global deformation well on average but may under-correct large local deformations near the diaphragm or lung periphery. A promising fix to this limitation is motion-weighted sampling, in which a coarse registration pass produces a rough displacement map that guides point selection toward high-motion regions for the final registration.

6. CONCLUSION

We have presented an adaptation of the Iterative Optimal Transport (IOT) framework to the problem of 3D intra-patient lung CT registration using pulmonary vessel point clouds. This project extended the original 2D multimodal formulation to a 3D point-cloud intra-patient formulation through three key modifications: a 3D Förstner corner detector for structure-aware point sampling, per-axis unit-variance normalization, and an evaluation protocol based on nearest-neighbor distance (motivated by the lack of manual landmark annotations).

Across eight experimental configurations on a COPD subject from the PVT1010 dataset, we find that the choice of point sampler has the greatest positive effect on registration performance. Förstner-scalar sampling with UOT-KL achieves a 59.9% reduction in mean nearest-neighbor distance (from 10.8 mm to 4.3 mm), compared with at most 7% for farthest-point sampling. Statistical testing confirms that Förstner significantly outperforms FPS in all matched comparisons ($p < 0.001$). While UOT-KL performs significantly better than UOT without KL in all configurations, increasing the polynomial degree from 2 to 3 yields only marginal, largely non-significant improvements.

These results demonstrate that the geometric focus of structure-aware sampling, combined with tolerance for mass imbalance via KL marginal penalties, is essential for effective 3D lung point cloud registration. Future work will extend evaluation to multiple subjects, incorporate motion-weighted sampling to reduce spatial variance, and benchmark against established registration baselines.

REFERENCES

[1] Li, M., Meng, C., and Fan, X., "Iterative optimal transport for multimodal image registration," *Pattern Recognition* **172**, 112736 (2026).

[2] Schöttle, F., Schneider, L., Hering, A., Tetzlaff, R., Kirchhoff, S., Maier-Hein, K. H., et al., "A combined 3d dataset for ct- and point cloud-based intra-patient lung registration," in [*Advances in Neural Information Processing Systems, Datasets and Benchmarks Track*], (2023). Lung250M-4B dataset.

[3] Haskins, G., Kruger, U., and Yan, P., "Deep learning in medical image registration: a survey," *Machine Vision and Applications* **31**(1), 8 (2020).

[4] Jiang, X., Ma, J., Xiao, G., Shao, Z., and Guo, X., "A review of multimodal image matching: methods and applications," *Information Fusion* **73**, 22–71 (2021).

[5] Haker, S., Zhu, L., Tannenbaum, A., and Angenent, S., "Optimal mass transport for registration and warping," *International Journal of Computer Vision* **60**(3), 225–240 (2004).

[6] Rehman, T. U., Haber, E., Pryor, G., Melonakos, J., and Tannenbaum, A., "3d nonrigid registration via optimal mass transport on the gpu," *Medical Image Analysis* **13**(6), 931–940 (2009).

[7] Feydy, J., Charlier, B., Vialard, F.-X., and Peyré, G., "Optimal transport for diffeomorphic registration," in [*Medical Image Computing and Computer-Assisted Intervention (MICCAI)*], *Lecture Notes in Computer Science* **10433**, 291–299, Springer (2017).

[8] Shen, Z., Feydy, J., Liu, P., Curiale, A. H., San José Estepar, R., San José Estepar, R., and Niethammer, M., "Accurate point cloud registration with robust optimal transport," in [*Advances in Neural Information Processing Systems*], **34**, 5373–5389 (2021).

[9] Motta, D., Casaca, W., and Paiva, A. L., "Vessel optimal transport for automated alignment of retinal fundus images," *IEEE Transactions on Image Processing* **28**(12), 6154–6168 (2019).

[10] Villani, C., [*Topics in Optimal Transportation*], vol. 58 of *Graduate Studies in Mathematics*, American Mathematical Society (2003).

[11] Ma, J., Jiang, X., Fan, A., Jiang, J., and Yan, J., "Image matching from handcrafted to deep features: A survey," *International Journal of Computer Vision* **129**, 23–79 (2021).

[12] Ma, J., Zhao, J., Ma, Y., and Tian, J., "Non-rigid visible and infrared face registration via regularized gaussian fields criterion," *Pattern Recognition* **48**(3), 772–784 (2015).

[13] Kim, S., Min, D., Ham, B., Do, M. N., and Sohn, K., "Dasc: Robust dense descriptor for multi-modal and multi-spectral correspondence estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39**(9), 1712–1729 (2017).

[14] Song, X., Xu, X., Zhang, J., Machado Reyes, D., and Yan, P., "Dino-reg: Efficient multimodal image registration with distilled features," *IEEE transactions on medical imaging* **44**, 3809–3819 (09 2025).

[15] Chizat, L., Peyré, G., Schmitzer, B., and Vialard, F.-X., "Scaling algorithms for unbalanced optimal transport problems," *Mathematics of Computation* **87**(314), 2563–2609 (2018).

[16] Chapel, L., Flamary, R., Wu, H., Févotte, C., and Gasso, G., "Unbalanced optimal transport through non-negative penalized linear regression," (2021).

[17] Förstner, W. and Gülch, E., "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in [*ISPRS Intercommission Workshop*], (June 1987).

[18] Villani, C., [*Topics in Optimal Transportation*], American Mathematical Society (2021). Originally published 2003.

[19] Shen, Z. et al., "Accurate point cloud registration with robust optimal transport." ArXiv.org (November 2021).