GaussianReg: Rapid 2D/3D Registration for Emergency Surgery via Explicit 3D Modeling with Gaussian Primitives

Supplementary Material

1. Arc-length Parameterization for Uniform Sampling

The ellipse is parameterized as

$$x = b^{'} cos\theta, \;\; y = c^{'} sin\theta, \tag{1}$$

the derivative $dx/d\theta$ and $dy/d\theta$ are

$$\frac{dx}{d\theta} = -b^{'} sin\theta, \;\; \frac{dy}{d\theta} = c^{'} cos\theta. \tag{2}$$

Hence, the differential arc length is given by

$$\frac{ds}{d\theta} = \sqrt{b'^2 sin^2\theta + c'^2 cos^2\theta}. \tag{3}$$

Then the differential value over the discretized angle steps can be computed by

$$ds = \sqrt{b'^2 sin^2\theta + c'^2 cos^2\theta} \times d\theta. \tag{4}$$

By integrating over the discretized angles, we form an approximate cumulative arc-length function

$$s(\theta^{'}) = \sum_{\theta=0}^{\theta^{'}} ds(\theta). \tag{5}$$

The cumulative arc-length is further normalized by $s_{norm}(\theta^{'}) = 2\pi \times s(\theta^{'})/P_{b^{'},c^{'}}$, so that it ranges from 0 to 2π. For the uniformly spaced target angles $\{\theta_k\}_{k=0}^{N_{point}-1}$ (with N_{point} sample points along the perimeter), we can obtain them by solving $s(\theta_k) = \frac{2\pi k}{N_{point}}$. In other words, the desired angles can be expressed as

$$\theta_k = s_{norm}^{-1}(\frac{2\pi k}{N_{point}}), k = 0, 1, \cdots, N_{point} - 1. \tag{6}$$

Once θ_k is determined, the spatial coordinate of the cell point to the ellipsoid center μ_i is calculated as $v_j^{'} = (-a + (m + 0.5)\frac{2a}{N_{ring}}, b^{'} cos\theta_k, c^{'} sin\theta_k)^T \in \mathbb{R}^3$.

2. More Quantitative Results

We show more quantitative results on the DeepFluoro dataset and Ljubljana dataset in Tab. 1 when only trained with 100 X-rays. As can be observed, even under time constraints and extremely limited data, our method still maintains registration accuracy close to 10 millimeters, demonstrating its robust capability.

Table 1. Comparison of our GaussianReg with different methods on the DeepFluoro dataset and Ljubljana dataset when training with 100 X-rays.

Method	DeepFluoro		Ljubljana	
	mTRE↓	mPD↓	mTRE↓	mPD↓
PoseNet [4]	32.94	24.94	62.58	44.58
Deep6DPose [1]	30.01	23.69	55.49	43.35
FixedLandmark [3]	24.03	18.48	–	–
GDR-Net [6]	20.93	16.00	38.26	29.04
RayEmb [5]	22.51	16.57	43.31	34.27
DiffPose [2]	20.90	15.47	39.04	29.86
Ours	**13.45**	**10.21**	**30.12**	**21.80**

Table 2. Ablation study on the number of rays Z sampled per peimitive and the number of rays k selected for pose estimation on the DeepFluoro dataset.

Method	mTRE ↓	mPD ↓
$Z = 50$	17.55	12.93
$Z = 100$	20.00	15.98
$\boldsymbol{Z = 200}$	**12.18**	**9.91**
$Z = 400$	14.89	11.41
$k = 25$	12.76	10.03
$\boldsymbol{k = 50}$	**12.18**	**9.91**
$k = 100$	15.19	11.95
$k = 150$	20.97	16.04

3. Extra Ablation Studies

We also conducted ablation study on the sampling number of rays Z per primitive and the selected number of rays k for direct pose estimation. Results are shown in Tab. 2. Increasing the number of rays Z enhances spatial resolution, enabling more precise camera search. However, excessive rays lead to increased computational load and reduced network efficiency, making it unsuitable for emergency scenarios. $Z = 200$ provides an excellent balance point. When k decreases, it means only the few rays most similar to the current X-ray are used for camera pose estimation; too small a number may compromise estimation stability. Conversely, when k increases, rays less similar to the X-ray are incorporated into pose estimation, introducing additional errors. By comparison, $k = 50$ proves optimal.

References

[1] Thanh-Toan Do, Ming Cai, Trung Pham, and Ian Reid. Deep-6dpose: Recovering 6d object pose from a single rgb image. *arXiv preprint arXiv:1802.10367*, 2018. 1

[2] Vivek Gopalakrishnan, Neel Dey, and Polina Golland. Intraoperative 2d/3d image registration via differentiable x-ray rendering. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 11662–11672, 2024. 1

[3] Robert B Grupp, Mathias Unberath, Cong Gao, Rachel A Hegeman, Ryan J Murphy, Clayton P Alexander, Yoshito Otake, Benjamin A McArthur, Mehran Armand, and Russell H Taylor. Automatic annotation of hip anatomy in fluoroscopy for robust and efficient 2d/3d registration. *Int. j. comput. assist. radiol. surg.*, 15:759–769, 2020. 1

[4] Alex Kendall, Matthew Grimes, and Roberto Cipolla. Posenet: A convolutional network for real-time 6-dof camera relocalization. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 2938–2946, 2015. 1

[5] Pragyan Shrestha, Chun Xie, Yuichi Yoshii, and Itaru Kitahara. Rayemb: Arbitrary landmark detection in x-ray images using ray embedding subspace. In *Asian Conf. Comput. Vis.*, pages 665–681, 2024. 1

[6] Gu Wang, Fabian Manhardt, Federico Tombari, and Xiangyang Ji. Gdr-net: Geometry-guided direct regression network for monocular 6d object pose estimation. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 16611–16621, 2021. 1