Embedded Feature Similarity Optimization with Specific Parameter Initialization for 2D/3D Medical Image Registration

Minheng Chen[1] Zhirun Zhang[1] Shuheng Gu[1] Youyong Kong[123]⋆

[1] School of Computer Science and Engineering, Southeast University, China
[2] Jiangsu Provincial Joint International Research Laboratory of Medical Information Processing, School of Computer Science and Engineering, Southeast University, China
[3] Key Laboratory of New Generation Artificial Intelligence Technology and Its Interdisciplinary Applications (Southeast University),Ministry of Education, Nanjing, China
⋆Corresponding author (Email: kongyouyong@seu.edu.cn)

INTRODUCTION

• Background

Image registration plays a crucial role in the field of medical image analysis and diagnosis.2D/3D registration is one of the most challenging problems in this field. This technique is primarily used for X-ray-based image-guided interventions and surgical image-based navigation, to estimate the spatial relationship between 3D preoperative CT and 2D intra-operative X-ray.

• Main challenges

1) Dimensional mismatch
2) Heavy computation
3) Lack of golden evaluation standard

• Motivation

1) Optimization-based methods are time consuming and limited by small capture range
2) Learning-based methods lack a method to initialize pose parameters
3) Current learning-based methods need large amounts of paired CTs and X-rays

• Innovation

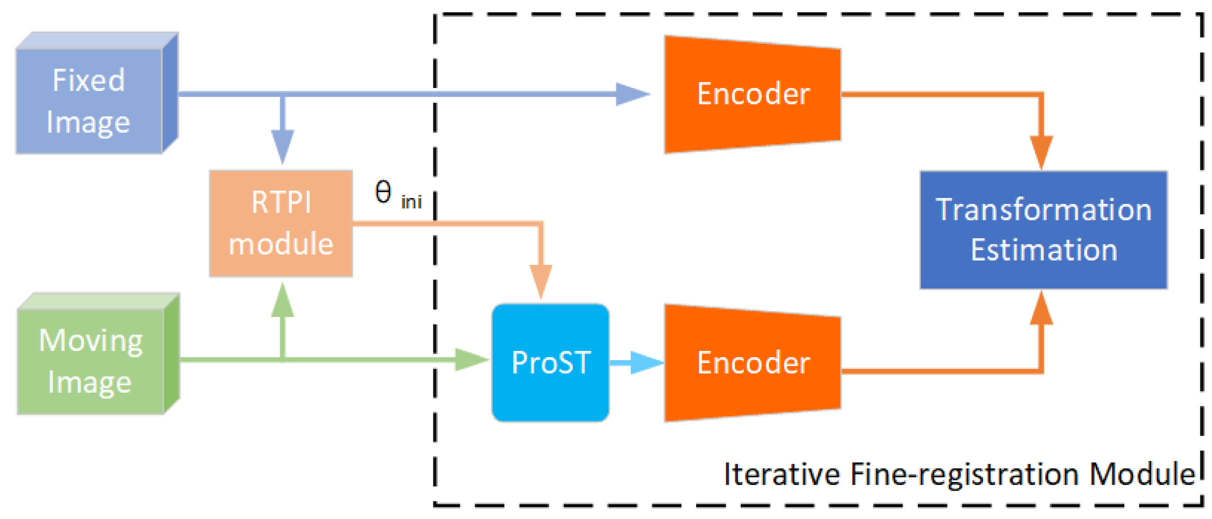

Fig. 1 The proposed two-stage 2D/3D registration framework

Methodology

• Problem definition

$$\mathcal{F}(\theta) = arg\min_{\theta} L_{sim}\left(I_f, P(\theta; V)\right)$$

The problem of 2D/3D registration is to seek a mapping function F to retrieve the pose parameter θ, where P(θ; V) denotes the mapping from volumetric 3D scene V to projective transmission image and θ is a 6 DoF vector (rx,ry,rz, tx, ty, tz).

• Rigid Transformation Parameter Initialization Module

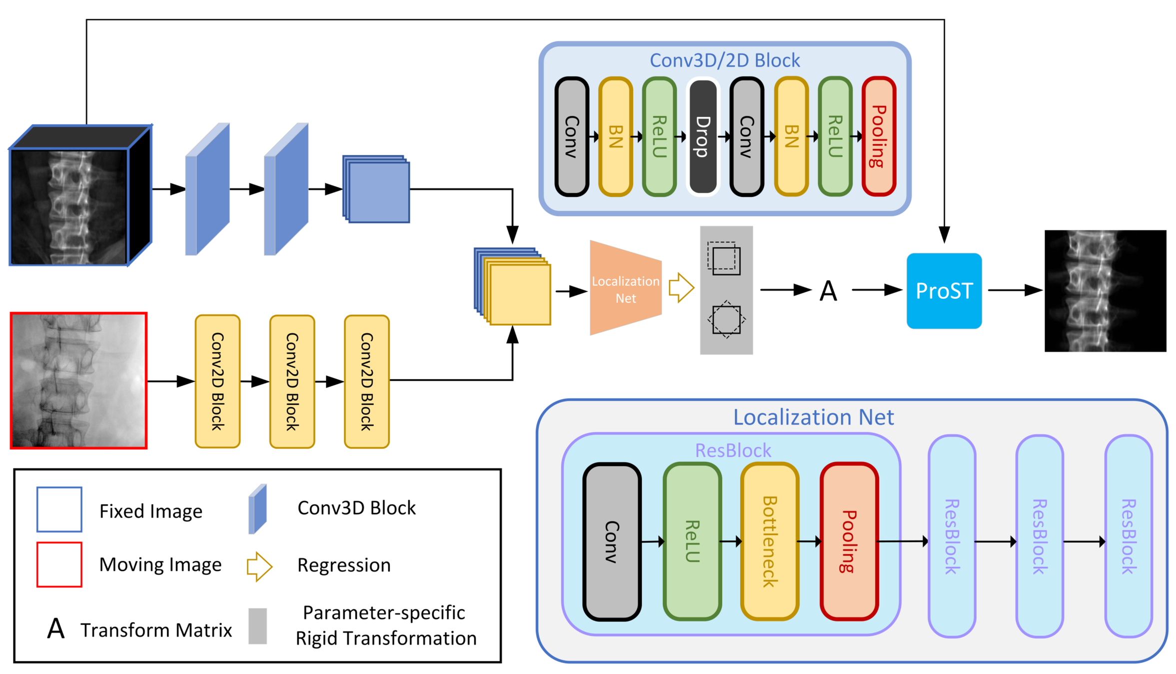

Fig. 2 The architecture of the Rigid Transformation Parameter Initialization module

To mitigate the dimension gap between 3D CTs and 2D X-rays

→ an asymmetric dual-branch structure to extract features

To ensure unique solutions and enhance interpretability

→ a parameter specific method to regress each transformation parameter

To train the network

$$\rightarrow \quad L_{mse}\left(\hat{\theta}, \theta\right) = \frac{1}{N}\sum_{i=1}^{N} |\hat{\theta}_i - \theta_i|_2$$

$$L_{RTPImodule} = \alpha L_{sim}\left(I_f, I_m\right) + \beta L_{mse}\left(\hat{\theta}, \theta\right) + \lambda \mathcal{R}(\theta)$$

L_{sim} is the gradient difference loss

• Iterative Fine-registration Module

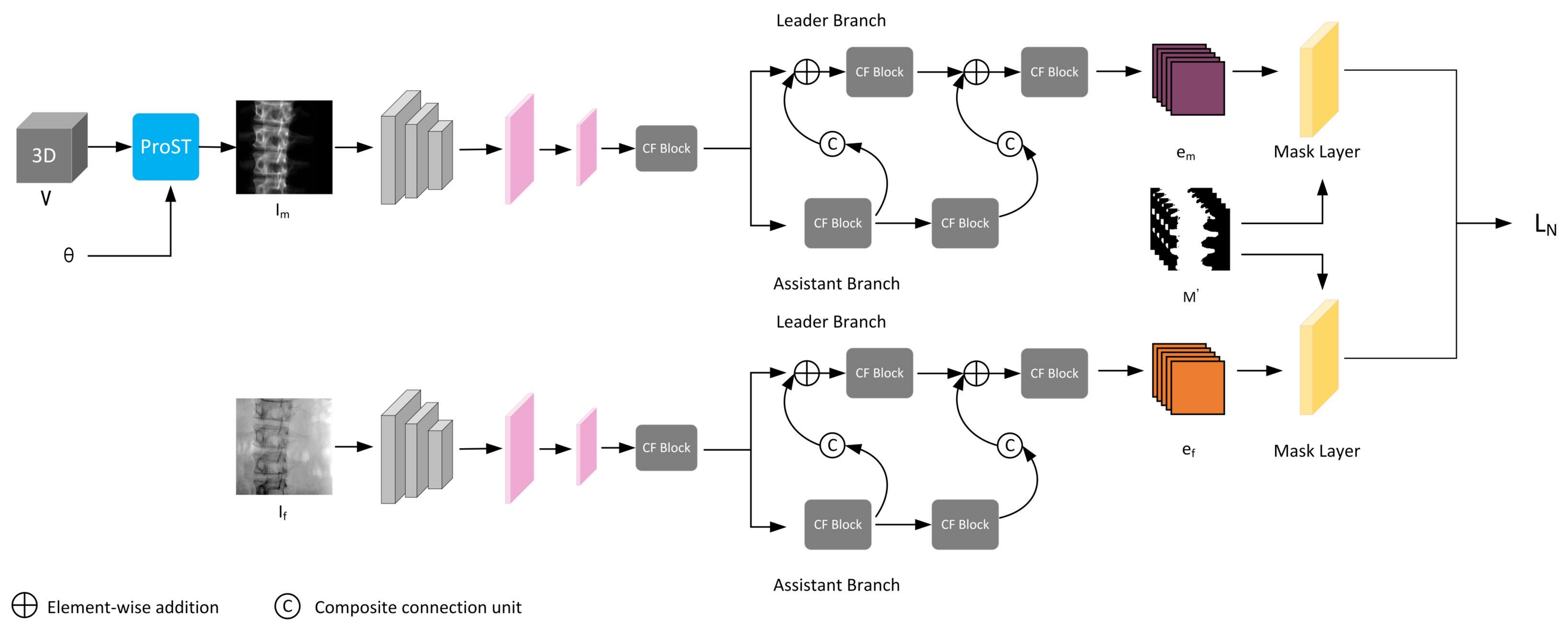

Fig. 3 The architecture of the Iterative Fine-registration module

Advantages: multi-scale feature fusion, expand capture range

the assistant branch extracts low-level features

fused to → the leader branch obtain both high and low level features

First back propagation: error function $L_N\left(e_m, e_f\right) = \frac{\sum_i^{H\times W\times C} M_i' \cdot |e_{m_i} - e_{f_i}|_2}{\sum_i^{H\times W\times C} M_i'}$

Second back propagation: real loss $L = |\frac{V_r'}{|V_r'|_2} - \frac{V_r}{|V_r|_2}|_2 + |\frac{V_t'}{|V_t'|_2} - \frac{V_t}{|V_t|_2}|_2$

Double backward mechanism

Experiment

• Dataset

146 raw CTS collected from hospitals, the spine is segmented using an automatic method and downsampled to the size 128×128×128. The simulation X-rays are generated by ProST module following a Perlove PLX118F C-Arm with settings that isotropic pixel is 0.19959 mm/pixel, the source-to-detector distance is 1011.7 mm and the detector dimension is 256×256.

• Qualitative result

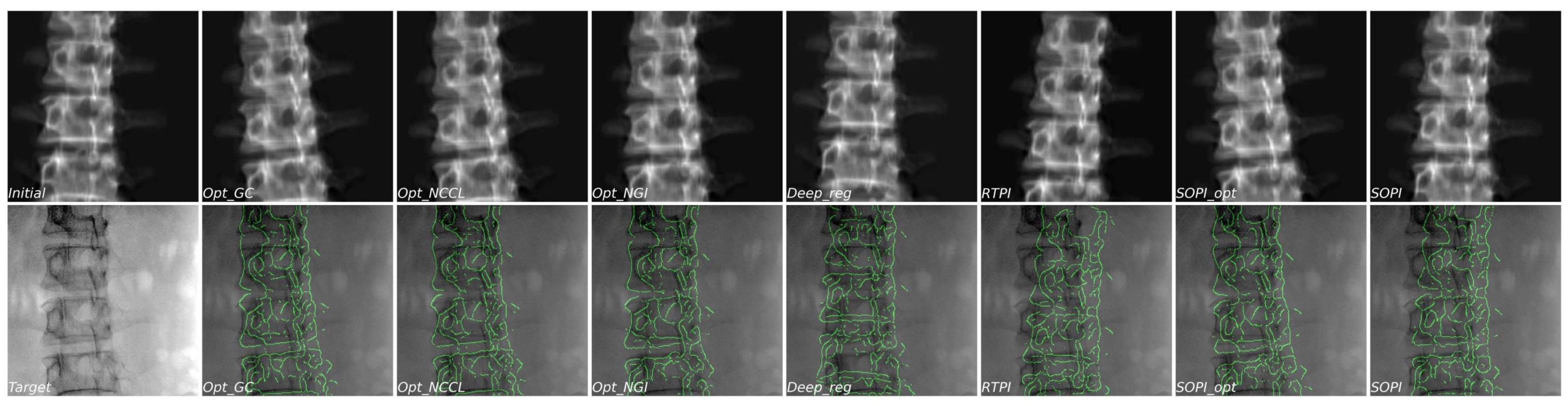

Fig. 4 Qualitative examples of our method and the baseline methods. The first row shows the projection results of the postures predicted by each method, and the second row shows the fusion images with the X-ray image, respectively.

• Evaluation and Results

We compare our method (SOPI) with one learning-based and four optimization-based methods. To further evaluate the performance of the proposed method as an initial pose estimator, we also demonstrate the performance of the method using our SOPI to initialize the optimization on X-ray data. We denote this approach as SOPI+opt.

Method	Rotation(°)	Translation(mm)	Failure rate(%)	Reg. time
Initial	6.40±3.77	15.08±8.56	95.2	N/A
Opt-NCCL [4]	3.68±3.18	5.79±5.18	38.2	19.16
Opt-NGI [5]	3.84±3.32	5.92±5.40	50.6	30.25
Opt-GC [6]	3.73±3.18	7.80±7.25	43.6	18.74
Deep-reg [15]	5.54±3.83	13.21±8.55	74.0	20.09
SOPI	**1.89±1.57**	**4.53±3.54**	**22.4**	**4.64**

Table 1: Performance comparison between our method and baseline methods on simulation.

Method	DisErr (mm)	ImgSim (NCC)	Rotation(°)			Translation(mm)			Reg. time
			rx	ry	rz	tx	ty	tz	
Initial	0.036	0.462	8.66	5.01	4.52	13.41	16.13	16.26	N/A
Opt-NCCL[4]	0.032	0.962	7.55	2.40	1.15	7.10	4.85	3.10	31.44
Opt-NGI[5]	0.036	0.941	7.75	5.00	1.30	18.15	4.80	**2.30**	48.64
Opt-GC[6]	0.027	0.939	4.70	2.40	1.60	7.15	3.15	3.00	40.14
Deep-reg[15]	**0.017**	**0.963**	6.80	7.93	4.07	5.82	6.67	10.08	20.08
SOPI	0.020	0.905	**2.83**	**0.90**	**0.88**	**5.64**	**1.50**	3.30	**4.67**
SOPI+opt	0.013	0.955	2.1	0.47	0.20	1.93	0.65	0.45	33.80

Table 2: Performance comparison between our method and baseline methods on X-ray. The performances are evaluated with distance error(DistErr), the image similarity score(ImgSim) and average registration time.

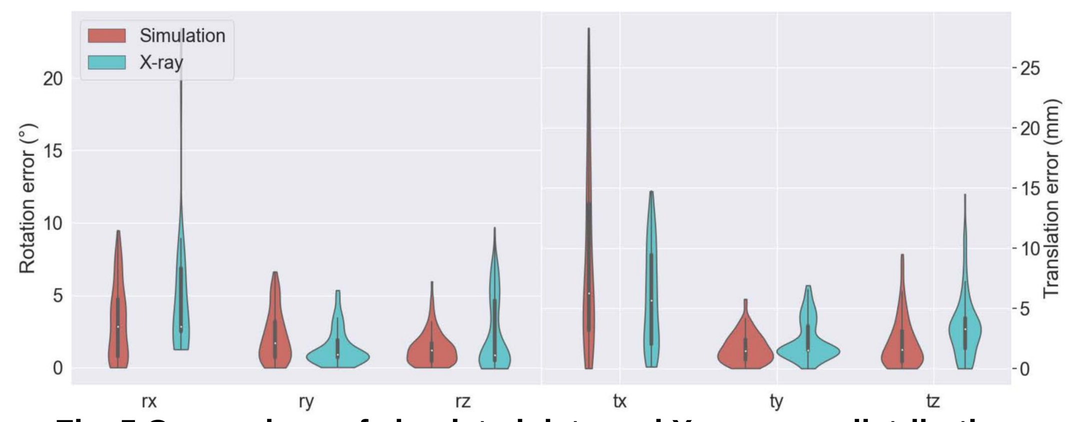

Fig. 5 Comparison of simulated data and X-ray error distribution.