

Non-Rigid Registration via Global to Local Transformation

Hao PAN, Yi MA, Fangrong ZHOU, Yan GU, Yutang MA, Chaobo MIN

Abstract: Non-rigid point set and image registration are key problems in plenty of computer vision and pattern recognition tasks. Typically, the non-rigid registration can be formulated as an optimization problem. However, registration accuracy is limited by local optimum. To solve this problem, we propose a method with global to local transformation for non-rigid point sets registration and it also can be used to infrared (IR) and visible (VIS) image registration. Firstly, an objective function based on Gaussian fields is designed to make a problem of non-rigid registration transform into an optimization problem. A global transformation model, which can describe the regular pattern of non-linear deformation between point sets, is then proposed to achieve coarse registration in global scale. Finally, with the results of coarse registration as initial value, a local transformation model is employed to implement fine registration by using local feature. Meanwhile, the optimal global and local transformation models estimated from edge points of IR and VIS image pairs are used to achieve non-rigid image registration. The qualitative and quantitative comparisons demonstrate that the proposed method has good performance under various types of distortions. Moreover, our method can also produce accurate results of IR and VIS image registration.

Keywords: Gaussian fields; global transformation; infrared image; non-rigid transformation; registration

1 INTRODUCTION

Image registration is a fundamental problem in many applications, such as pattern recognition, image mosaic and stereo vision, especially multi-sensor image fusion [1-4]. These applications must be on the basis of successful image registration. The complementary information from multi-sensor can offer more varied and comprehensive scene representations. It is very helpful for human visual perception, object detection, tracking and recognition. Multi-sensor image registration is a prerequisite for image fusion. Therefore, it has always been a research hotspot [5-10].

However, due to the complementary of multi-sensor images, there is less mutual information between them, such as infrared (IR) and visible (VIS) images. Some features in VIS images are commonly loss in IR images. Thus, multi-sensor image registration, which is very critical for fusion, is still a challenging task. One of the solutions is to transform image registration into point set registration, and then estimate spatial transformation model from point feature [5, 11]. This paper focuses on point set registration to achieve IR and VIS image registration.

Spatial Transformation model is very important for registration. Moreover, it directly determines registration accuracy and computational complexity. At present, various transformation models are proposed for point set and image registration. Overall, the common used transformation models include rigid model, affine model [12], thin-plate spline (TPS) model [13], B-spline model [14] and the transformation model within a vector-valued reproducing kernel Hilbert space (RKHS) [15].

Rigid transformation contains only translation and rotation. On the basis of rigid transformation, the affine transformation contains scaling and skews additionally. The affine model has more dimensions of transformation than rigid model and is suitable for non-rigid registration. Because of its simplicity and high performance, affine transformation has been widely used. The key ideal is to estimate the coefficients of the affine model from correspondence between point sets. In [16], the correspondence point pairs, which are extracted from IR

and VIS images by visual saliency and optimization procedure based on random sample consensus (RANSAC), are applied to calculate the coefficients of the affine transformation model. In [17], the optimal affine model is estimated from the correspondence obtained by ranking the geometric similarity and Euler distance. In [18], silhouette extraction is introduced to select the matching point pairs between IR and VIS images so that the affine transformation coefficients can be determined. However, the affine transformation is linear and cannot produce accurate alignment when there exists the anisotropy of deformation between point pairs in many applications, especially in multi-sensor image fusion [19, 20]. Therefore, some non-linear transformation models have been proposed to address this problem.

TPS is a good non-linear and non-rigid transformation model. It can be cleanly decomposed into affine and non-affine subspaces while minimizing a bending energy, and can be regarded as a non-rigid extension of affine transformation. Because the transformation coefficients of the TPS are more than that of affine model, the algorithms used to estimate the optimal TPS model are more complex. In [21], a robust point matching (RPM) is developed with the TPS as the parameterization of the non-rigid transformation and the soft assign of the correspondence. In [22], Speed-up robust feature (SURF) is introduced to extract corresponding feature point pairs so that the optimal TPS coefficients are determined. Changcai Yang et al. [23] propose a self-adaptive weighted objective function based on mixture model, and use expectation-maximization (EM) algorithm as an optimization procedure to achieve non-rigid point set registration.

Because B-spline is able to preserve the smoothness of non-linear deformation, it is commonly utilized to construct non-linear spatial transformation model. Suicheng Gu et al. [24] propose a B-spline affine transformation and used the iterative closest point (ICP) method to achieve the registration of three-dimensional (3D) volumetric computed tomography (CT) data. In [25], a B-spline transformation model is constructed by control point parameterization and a gradient optimization procedure is applied to obtain the optimal B-spline basis coefficients. Wei Sun et al. [25] utilize lower-order B-

spline basis functions and a random perturbation technique to implement efficient non-rigid registration.

Jiayi Ma et al. [15] propose a non-linear transformation model within RKHS. With the RKHS-based transformation model, the displacement of points is determined by a high dimensional coefficient vector constructed from control points in neighborhood. Due to high flexibility of transformation, it has been applied to various registration approaches. In [27], the L_2 -minimizing estimator is presented to estimate the RKHS-based transformation coefficients by building robust sparse and dense correspondences between point sets. Ref. [28] introduces an objective function established by Gaussian Mixture Models (GMMs) to measure the registration performance under different transformation coefficients and EM algorithm is employed to determine the global optimal solution.

The non-linear transformation models used in existing registration methods are based on the assumption that the displacement of every point largely depends on local feature in neighborhood. They describe non-linear deformation between point pairs by using feature points or control points. However, the optimization of registration may be accompanied by local optimum and premature convergence. Aimed at this problem, this paper proposes a method with global to local transformation for non-rigid point set and image registration. An objective function is established with Gaussian fields to make a problem of non-rigid registration transform into an optimization problem. The registration in this paper proceeds in two stages called coarse registration and fine registration respectively, in order to overcome local convergence. At the stage of coarse registration, a global transformation model, which can describe the regular pattern of non-linear deformation between all point pairs in global scale, is designed and applied to achieve coarse registration. With the result of coarse registration as initial value, a local transformation model is then employed to complete final registration at the stage of fine registration. To achieve non-rigid image registration, the edge maps are extracted from IR and VIS images by the Canny edge detector, and then, the optimal image transformation model is estimated from edge maps by using the strategy with global to local transformation.

The primary contributions of this work are as follows:

(1) A global transformation model, which is able to describe the regular pattern of global deformation between point sets, is proposed to achieve coarse registration to overcome local convergence.

(2) A strategy with global to local transformation is designed and applied to improve the accuracy of non-rigid registration.

The rest of the paper is organized as follows. In section 2, we present the details of the proposed method. Section 3 illustrates the proposed method on synthesized point set registration, and then tests it on real images under various scenarios with comparisons to other approaches. Finally, section 4 presents the concluding remarks for our work.

2 PROPOSED METHOD

Essentially, our method is a gradient-based optimization process for non-rigid registration. So this section mainly focuses on objective function, global

transformation, local transformation and optimization procedure.

2.1 Objective Function for Non-Rigid Registration

$U = \{\mathbf{u}_m\}_{m=1}^M$ and $V = \{\mathbf{v}_n\}_{n=1}^N$ ($\mathbf{u}_m, \mathbf{v}_n \in \mathcal{R}^2$) are used to represent two point sets, where U is considered as ‘model’ point set and V is ‘data’ point set. The purpose of registration is to align the model point set U to the data point set V . We represent the non-rigid transformation by a map function f . A point \mathbf{u}_m is mapped to a new location $\hat{\mathbf{u}}_m = f(\mathbf{u}_m)$. \hat{U} is the whole transformation point set U , and $\hat{U} = \{\hat{\mathbf{u}}_m\}_{m=1}^M$.

The objective function for registration is essentially a criterion that can quantify the performance of registration with much accuracy. In this paper, Gaussian fields are used to establish the objective function as follow,

$$\begin{aligned} \min_f E(f) &= \\ &= \min_f - \sum_{m=1}^M \sum_{n=1}^N c(\mathbf{v}_n, f(\mathbf{u}_m)) \exp\left(-\frac{\|\mathbf{v}_n - f(\mathbf{u}_m)\|^2}{2\sigma_d^2}\right) + \lambda S(f) \end{aligned} \quad (1)$$

where $c(\mathbf{v}_n, f(\mathbf{u}_m))$ denotes the correspondence between \mathbf{v}_n and $f(\mathbf{u}_m)$, $\|\cdot\|$ denotes the L_2 norm, σ_d is a range parameter. The first term of Eq. (1) is used to measure the Euclidean distance between corresponding points. The second term ensures that the objective function is a smooth optimization problem, enforcing smoothness to the non-rigid transformation function f . λ is a regularization constant that balances these two terms. The optimal transformation f is determined by minimizing the objective Eq. (1).

Let C represent the correspondence matrix between U and V , and $C_{mn} = c(\mathbf{v}_n, f(\mathbf{u}_m))$. Ideally, C is supposed to be a binary matrix. If \mathbf{v}_n is corresponding to \mathbf{u}_m , $C_{mn} = 1$, otherwise $C_{mn} = 0$. However, it is very difficult to obtain a real corresponding matrix unless manual annotation is used in general. Hence, we need to construct a model which is able to indicate the potential correspondence between two point sets. In this paper, C_{mn} is defined as

$$C_{mn} = c(\mathbf{v}_n, f(\mathbf{u}_m)) = \frac{\exp\left(-\frac{R_{mn}^2}{2\sigma_s^2} - \frac{\|\mathbf{v}_n - f(\mathbf{u}_m)\|^2}{2\sigma_d^2}\right)}{\sum_{a=1}^M \exp\left(-\frac{R_{an}^2}{2\sigma_s^2}\right) + \sum_{b=1}^N \exp\left(-\frac{R_{mb}^2}{2\sigma_s^2}\right)} \quad (2)$$

where R is an attribute matrix which indicates similarity between U and V . In Eq. (2), the row and column summation constraints guarantee that the correspondence is one-to-one. Because this work focuses on 2D point registration, shape context [29] is used to construct the attribute matrix R . R_{mn} can be written as

$$R_{mn} = \frac{1}{2} \sum_{i=1}^I \frac{(S_i(\mathbf{u}_m) - S_i(\mathbf{v}_n))^2}{S_i(\mathbf{u}_m) + S_i(\mathbf{v}_n)} \quad (3)$$

where $S_i(\mathbf{u}_m)$ and $S_i(\mathbf{v}_n)$ denote the I -bin normalized histogram at \mathbf{u}_m and \mathbf{v}_n , respectively. Substituting Eq. (2) into Eq. (1), the objective function becomes

$$\min_f E(\mathbf{f}) = \min_f \sum_{m=1}^M \sum_{n=1}^N w_{mn} \exp\left(-\frac{\|\mathbf{v}_n - \mathbf{f}(\mathbf{u}_m)\|^2}{\sigma_d^2}\right) + \lambda S(\mathbf{f}) \quad (4)$$

where w_{mn} is defined as

$$w_{mn} = \exp\left(-\frac{R_{mn}^2}{2\sigma_s^2}\right) \left/ \left(\sum_{a=1}^M \exp\left(-\frac{R_{an}^2}{2\sigma_s^2}\right) + \sum_{b=1}^N \exp\left(-\frac{R_{mb}^2}{2\sigma_s^2}\right) \right) \right. \quad (5)$$

2.2 Global Transformation Model

Global transformation model is represented as f_G . Let $\mathbf{u}_m = [x, y]$ and $f_G(\mathbf{u}_m) = [\hat{x}, \hat{y}]$. f_G can be written as

$$\begin{cases} \hat{x} = [x, y, 1] \times [s_x \cos(\theta), -\sin(\theta), t_x]^T + G(x, y, \mathbf{K}) \\ \hat{y} = [x, y, 1] \times [\sin(\theta), s_y \sin(\theta), t_y]^T + G(x, y, \mathbf{P}) \end{cases} \quad (6)$$

where θ is the angle of rotation in affine transformation, s_x and s_y are the scaling coefficients, t_x and t_y are the translation coefficients. $G(x, y, \zeta)$ represents the model of non-linear transformation and it is formulated as

$$G(x, y, \zeta) = \sigma \sum_{h=2}^4 \sum_{g=0}^h \zeta_{h,g} x^g y^{h-g} \quad (7)$$

From Eq. (7) we can see that the model of non-linear transformation consists of quadratic, cubic and quartic polynomial. ζ is a coefficient vector of polynomial transformation. σ is a constant that balances affine and polynomial transformation models. We set $\sigma = 1 \times 10^{-4}$ throughout this work.

Substituting Eq. (7) into Eq. (6), the matrix form of the global transformation model becomes

$$\mathbf{f}_G(\mathbf{u}_m) = \mathbf{G}_m [\mathbf{A} | [\mathbf{K} | \mathbf{P}]^T]^T \quad (8)$$

where global transformation matrix $\mathbf{G}_m = [x, y, 1, \sigma x^4, \sigma y^4, \sigma x y^3, \sigma x^3 y, \sigma x^2 y^2, \sigma x^2, \sigma y^2, \sigma x y^2, \sigma x^2 y, \sigma x^2, \sigma y^2, \sigma x y]$, $[\mathbf{A} | [\mathbf{K} | \mathbf{P}]^T]^T$ is called as coefficient matrix of global transformation, $\mathbf{A} = [s_x \cos(\theta), -\sin(\theta), t_x; \sin(\theta), s_y \cos(\theta), t_y]$, $\mathbf{K} = [k_{4,0}, k_{4,4}, k_{4,1}, k_{4,3}, k_{4,2}, k_{3,0}, k_{3,3}, k_{3,1}, k_{3,2}, k_{2,0}, k_{2,2}, k_{2,1}]^T$, $\mathbf{P} = [p_{4,0}, p_{4,4}, p_{4,1}, p_{4,3}, p_{4,2}, p_{3,0}, p_{3,3}, p_{3,1}, p_{3,2}, p_{2,0}, p_{2,2}, p_{2,1}]^T$. \mathbf{K} and \mathbf{P} are coefficient vectors of the polynomial transformation in the x and y directions, respectively. Both of them are 12×1 dimensional vectors.

Essentially, the global transformation model consists of two parts: affine transformation and polynomial transformation. Affine transformation model is constructed by \mathbf{A} and the first three elements of \mathbf{G}_m . The polynomial transformation model, which is employed to describe the regular pattern of non-linear deformation between point sets in this work, is formed with \mathbf{K} , \mathbf{P} and the rest of \mathbf{G}_m . In the global transformation model, the displacements between all point pairs exhibit a non-linear regular pattern based on polynomial. The coefficients of affine and non-

affine transformations can be optimized within a uniform pattern so that the global convergence could balance against local convergence.

The coefficient vector of global transformation, which contains affine and polynomial transformation coefficients, is defined as $\mathbf{Q} = [\theta, s_x, s_y, t_x, t_y | \mathbf{K}^T | \mathbf{P}^T]^T$, while it is a 29×1 dimensional vector. Substituting Eq. (8) into the objective function (4), the objective function of the coarse registration becomes

$$\min_{\mathbf{Q}} E(\mathbf{Q}) = \min_{\mathbf{Q}} \sum_{m=1}^M \sum_{n=1}^N w_{mn} \exp\left(-\frac{\|\mathbf{v}_n - \mathbf{H}_m\|^2}{\sigma_d^2}\right) + \lambda \cdot \text{tr}((\mathbf{Q} - \mathbf{Z})(\mathbf{Q} - \mathbf{Z})^T) \quad (9)$$

where $\mathbf{H}_m = \mathbf{G}_m [\mathbf{A} | [\mathbf{K} | \mathbf{P}]^T]^T$, $\mathbf{Z} = [0, 1, 1, 0, \dots]^T$ is a 29×1 dimensional vector, $\text{tr}(\cdot)$ denotes the trace.

2.3 Local Transformation Model

In this work, the RKHS-based transformation in [15] is employed as the local transformation model. The RKHS-based transformation model is chosen because it can be optimized by using local neighborhood structure. The local transformation model f_L is defined as

$$[\hat{x}, \hat{y}]^T = [x, y]^T + \mathbf{f}_L(\mathbf{u}_m) = [x, y]^T + \sum_{i=1}^M \Phi(\mathbf{u}_m, \mathbf{u}_i) \mathbf{t}_i \quad (10)$$

where a diagonal Gaussian kernel

$\Phi(\mathbf{u}_i, \mathbf{u}_j) = \exp(-\beta \|\mathbf{u}_i - \mathbf{u}_j\|^2) \cdot \mathbf{I}$, $\mathbf{I} = [1, 0; 0, 1]$, \mathbf{t}_i is a 2×1 dimensional coefficient vector of local transformation in feature point \mathbf{u}_i .

Substituting Eq. (10) into the objective function (4), the objective function of the fine registration becomes

$$\min_{\mathbf{T}} E(\mathbf{T}) = \min_{\mathbf{T}} \sum_{m=1}^M \sum_{n=1}^N w_{mn} \exp\left(-\frac{\|\mathbf{v}_n - \mathbf{u}_m - \Phi_m \mathbf{T}\|^2}{\sigma_d^2}\right) + \lambda \cdot \text{tr}(\mathbf{T}^T \Phi \mathbf{T}) \quad (11)$$

where $\mathbf{T} = [t_1, \dots, t_M]^T$ is the coefficient matrix of size $M \times 2$, the kernel matrix $\Phi \in \mathfrak{R}^{M \times M}$ is called the Gram matrix with $\Phi_{ij} = \exp(-\beta \|\mathbf{u}_i - \mathbf{u}_j\|^2)$, Φ_m denotes the m^{th} row of Φ .

2.4 Optimization

The objective function (9) and (11) are based on Gaussian fields. Thus, they are continuously differentiable with respect to the transformation coefficient \mathbf{Q} and \mathbf{T} , respectively. The derivative of Eq. (9) is given by

$$\frac{\partial E(\mathbf{Q})}{\partial \mathbf{Q}} = \sum_{m=1}^M \sum_{n=1}^N \frac{2w_{mn}}{\sigma_d^2} \cdot \frac{\partial \mathbf{H}_m}{\partial \mathbf{Q}} (\mathbf{H}_m - \mathbf{v}_n)^\top \cdot \exp\left(-\frac{\|\mathbf{v}_n - \mathbf{H}_m\|^2}{\sigma_d^2}\right) + 2\lambda \cdot (\mathbf{Q} - \mathbf{Z}) \quad (12)$$

where $\partial \mathbf{H}_m / \partial \mathbf{Q}$ is 29×2 dimensional matrix and is obtained by

$$\frac{\partial \mathbf{H}_m}{\partial \mathbf{Q}} = \begin{bmatrix} -s_x x \sin \theta - y \cos \theta & x \cos \theta - s_y y \sin \theta \\ x \cos \theta & 0 \\ 0 & y \cos \theta \\ 1 & 0 \\ 0 & 1 \\ \mathbf{G}_m^s & 0 \\ 0 & \mathbf{G}_m^s \end{bmatrix} \quad (13)$$

where $\mathbf{G}_m^s = [\alpha x^4, \sigma y^4, \alpha x y^3, \alpha x^3 y, \alpha x^2 y^2, \alpha x^3, \sigma y^3, \alpha x y^2, \alpha x^2 y, \alpha x^2, \sigma y^2, \alpha x y]^\top$

The derivative of Eq. (11) can be computed using

$$\frac{\partial E(\mathbf{T})}{\partial \mathbf{T}} = \sum_{m=1}^M \sum_{n=1}^N \frac{2w_{mn}}{\sigma_d^2} \cdot \Phi_m^\top (\Phi_m \mathbf{T} + \mathbf{u}_m - \mathbf{v}_n) \cdot \exp\left(-\frac{\|\mathbf{v}_n - \mathbf{u}_m - \Phi_m \mathbf{T}\|^2}{\sigma_d^2}\right) + 2\lambda \cdot \Phi \mathbf{T} \quad (14)$$

Algorithm 1 Non-rigid registration with global to local transformation

Input: Two point sets $\mathbf{U} = \{\mathbf{u}_m\}_{m=1}^M$ and $\mathbf{V} = \{\mathbf{v}_n\}_{n=1}^N$, corresponding attribute sets $\{S(\mathbf{u}_m)\}_{m=1}^M$ and $\{S(\mathbf{v}_n)\}_{n=1}^N$, parameters β , λ , σ_d and σ_s .

Output: The optimal transformation \mathbf{F}

- 1 Construct the attribute matrix \mathbf{R} ;
- Coarse registration:
- 2 Initialize the coefficient vector \mathbf{Q} of the global transformation;
- 3 By using the derivative (12), optimize the objective function (9) by quasi-Newton method, and then obtain the optimal transformation coefficient vector \mathbf{Q} ;
- 4 Compute the transformation result $f_G(\mathbf{U})$ by using Eq. (8);
- Fine registration:
- 5 With $f_G(\mathbf{U})$ as input, initialize the coefficient vector \mathbf{T} of the local transformation;
- 6 By using the derivative (14), optimize the objective function (11) by quasi-Newton method, and then obtain the optimal transformation coefficient matrix \mathbf{T} ;
- 7 Compute the final transformation result $\mathbf{F} = f_i(f_G(\mathbf{U}))$ by using Eq. (10).

3 EXPERIMENT

To appraise our method, we first test it on synthesized point sets, and then focus on registration for real IR and VIS images. Our method requires four parameters to be set: β , λ , σ_d and σ_s . We set $\beta = 0.2$, $\lambda = 0.1$, $\sigma_d = 5$ and $\sigma_s = 0.3$ throughout this work.

The algorithms are implemented in Matlab and run on the computer with 3.9 GHz Intel Core CPU, 4 GB memory. The average runtime when using the proposed method for point set registration is about 35 seconds, and the average runtime of IR and VIS image registration without edge discretization is about 1 minute.

With Eq. (12) and Eq. (14), a gradient-based numerical optimization technique can be used to determine the optimal transformation coefficient. In this paper, quasi-Newton method is employed to solve the optimization problem. However, quasi-Newton method has high requirement for initial value, otherwise the optimization procedure is limited by local convergence. Thus, a coarse-to-fine strategy with global to local transformation is designed to improve the chance of reaching the global minimum. At the stage of coarse registration, the corresponding objective function is minimized by using the global transformation model. At the stage of fine registration, with the result of coarse registration as the initial value, the final registration is then achieved by using the objective function of fine registration with the local transformation model. Because the global transformation model contains affine and polynomial transformation, the offsets of all points are conformed to a regular pattern. Thus, the coarse registration focuses on optimal transformation in global scale. Due to the local transformation model constructed with local feature, individual offset of every point is optimized in its neighborhood at the stage of fine registration. The initial data estimated by the coarse registration is used to improve the performance of the fine registration. The proposed method is outlined in Algorithm 1.

3.1 Results on Point Set Registration

The synthesized data sets are constructed by Chui and Rangarajan [21], and they consist of two different point sets which are respectively named by *Fish* and *Fu* in this paper. The point set of *Fish* contains 98 points and the point set of *Fu* contains 105 points, as shown in Fig. 1.

Fig. 1 reports the results of registration with the global, local and global to local transformation models, respectively. We see that the results of registration with the global transformation or the local transformation are not accurate enough. It illustrates that the optimization with the global transformation or the local transformation is easy to fall into local convergence and does not reach global minimum. Meanwhile, compared to the results of registration with the global transformation, the

performance of registration with the local transformation is degraded more easily by local convergence, especially for *Fish* point sets. This illustrates that the global transformation model used in the coarse registration is able to prevent optimization from falling into local

convergence. Apparently, our method can produce more accurate alignment results. It demonstrates that a good initial data estimated by the coarse registration improves the performance of the fine registration.

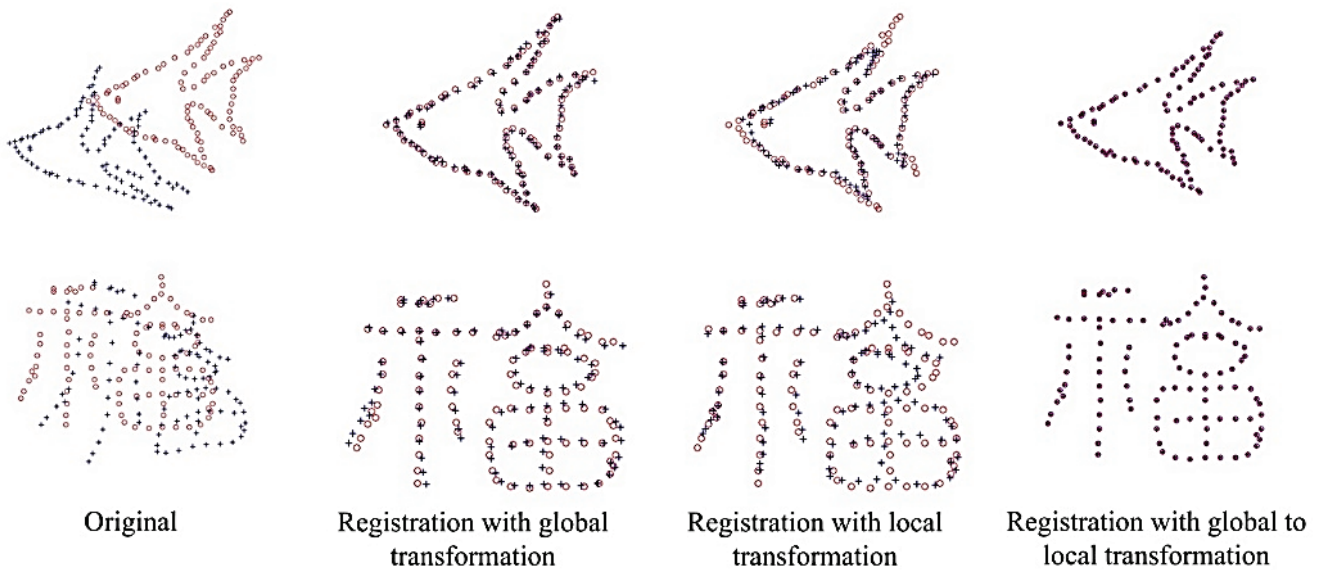


Figure 1 The results of registration with the global, local and global to local transformation models

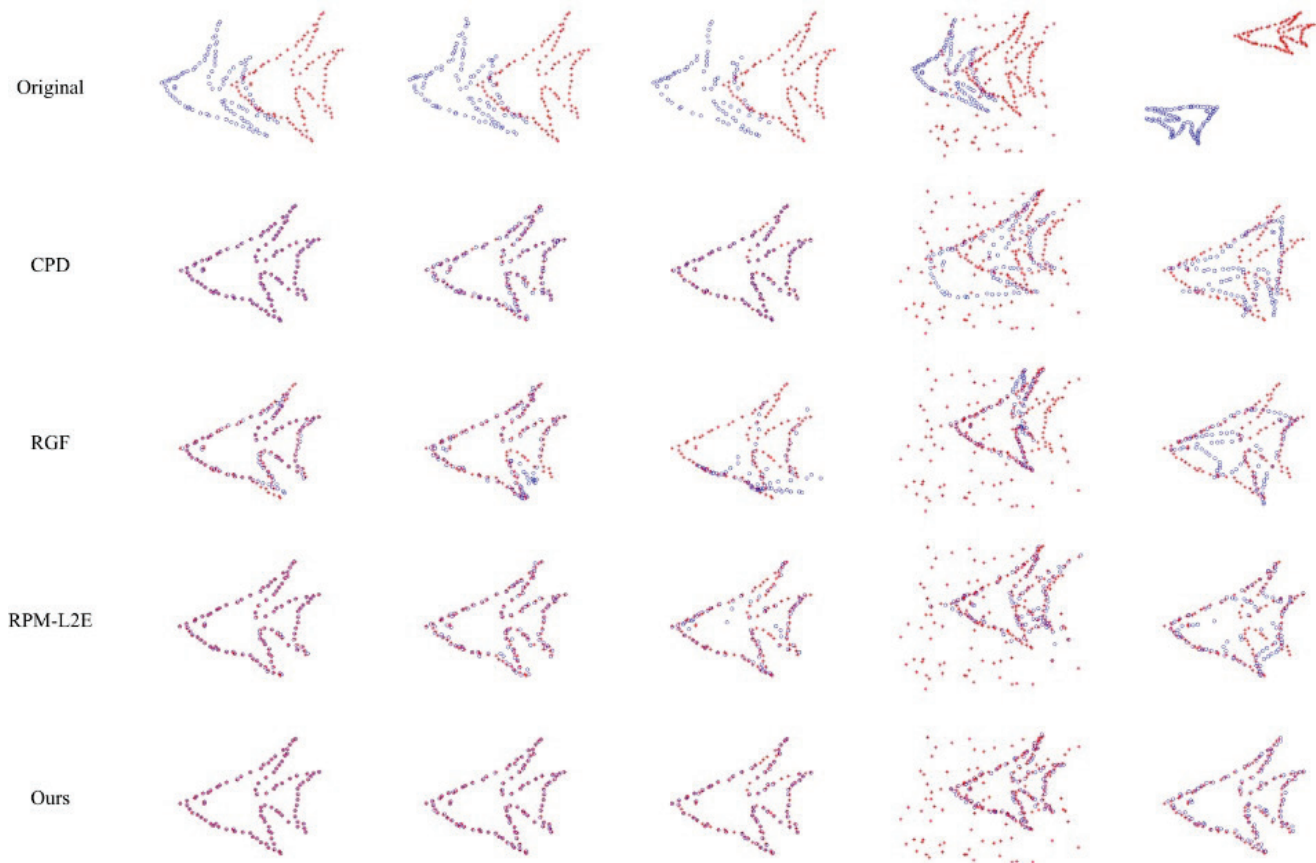


Figure 2 The examples of registration results by the various methods on *Fish* point sets degraded by the degeneration models of deformation, noise, outlier, occlusion and rotation. Top row: the model (blue 'o') and target (red '*') point sets. From the second row and the bottom row: the registration results by CPD, RGF, RPM-L₂E and the proposed method, respectively.

To evaluate the proposed method, we test it compared to CPD [32], RGF [15] and RPM-L₂E [27], which are the state-of-the-art non-rigid registration methods. The Matlab codes of the above algorithms are provided by their authors.

The five models of degeneration including deformation, noise, occlusion, outlier and rotation are employed to test the performance of the above non-rigid registration algorithms. In order to analyze the influence of point set

distortion, each above degeneration model is used to generate a target set from an individual point set, and 100 samples are created for each degradation level. Some

examples from Fish and Fu point sets are shown in the top row of Fig. 2 and Fig. 3.

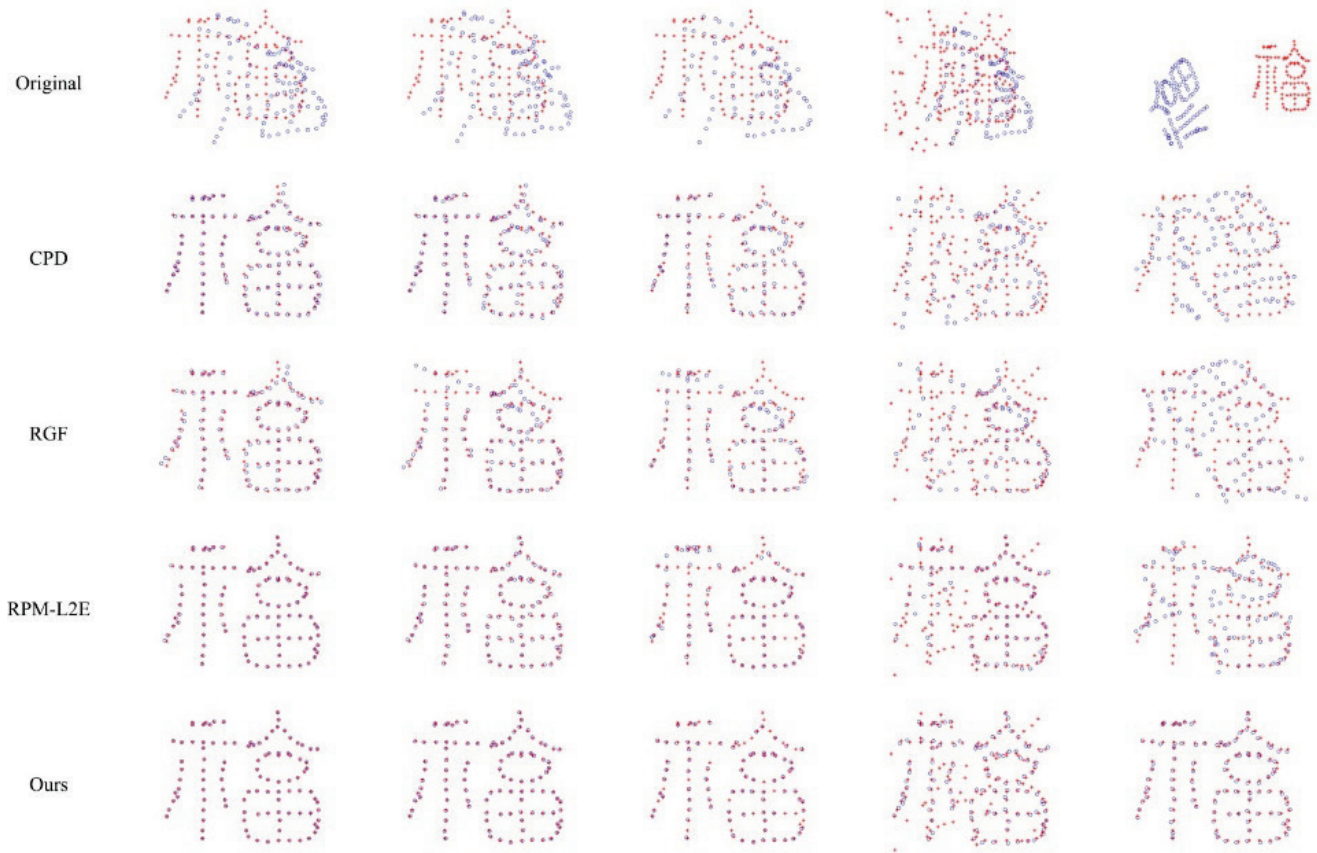


Figure 3 The examples of registration results by the various methods on *Fu* point sets degraded by the degeneration models of deformation, noise, outlier, occlusion and rotation. Top row: the model (blue 'o') and target (red '*') point sets. From the second row and the bottom row: the registration results by CPD, RGF, RPM-L₂E and the proposed method, respectively.

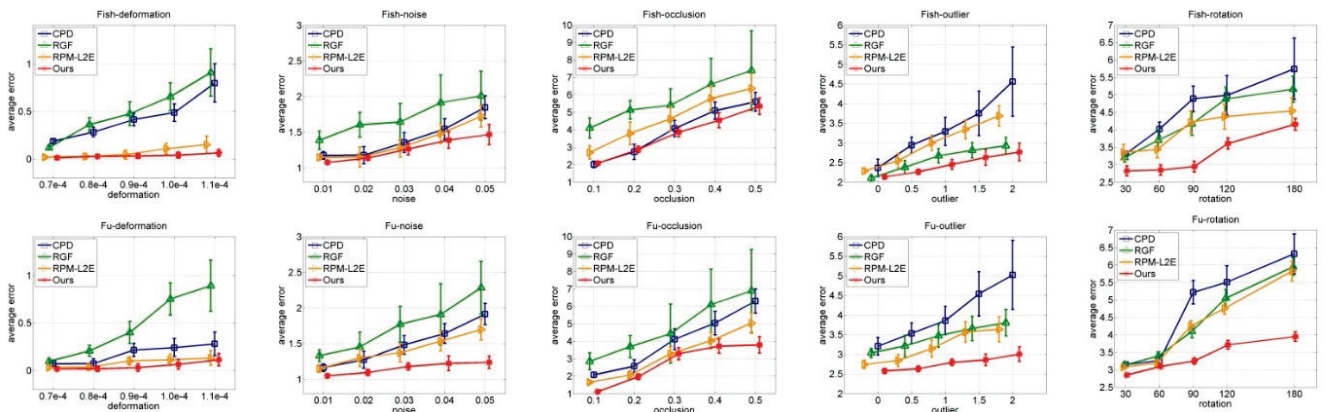


Figure 4 Quantitative comparison of our method with CPD, RGF and RPM-L₂E. The error bars indicate the registration error means and standard deviations over 100 trials.

The qualitative results achieved by the various registration methods are also shown in Fig. 2 and Fig. 3. CPD performs well on point registration under the degeneration models of deformation, noise and occlusion, especially for Fish point set, while it is failed to implement point registration under the degeneration models of outline and rotation.

RGF is constantly easy to fall into local convergence and is not able to achieve point registration very well. RPM-L₂E has a good performance under the degeneration models of deformation, noise, occlusion and outline,

especially for *Fu* point set. However, it is not able to deal with the degeneration model of rotation adequately. Compared with CPD, RGF and RPM-L₂E, the proposed method produces the best results of point registration under the various degeneration models. It can be seen that the proposed method with global to local transformation has a good chance of reaching the global optima.

The root-mean-square error (RMSE) is used for quantitative evaluation of the various registration results on the synthesized point sets, where $RMSE = \sqrt{\sum \|\hat{U} - V\|^2 / M}$.

Fig. 4 reports the quantitative results of the registration methods. In most cases, our method has the lowest registration error mean and standard deviation compared with the other approaches. The advantage of our method becomes more and more obvious as the degree level of degradations increases. This is because the proposed method with global to local optimization helps to estimate non-rigid transformation robustly.

3.2 Result on IR and VIS Image Registration

In this section, our method is tested on the four pairs of IR and VIS image pairs, i.e. *Stairs*, *Building*, *Car* and *Block*, which are from CVC datasets [30]. The resolutions of all images are 320×240 . Meanwhile, the performance of our method is compared with those of the state-of-the-art non-rigid registration methods: CPD, RGF and RPM-L2E. All of these methods are able to estimate transformation model from point feature.

The edge maps are obtained by Canny edge detector, and a sampling method introduced in [29] is employed to discretize the edge maps into a point set. The goal is then to register the IR images to the VIS images. To present registration results visually, a simple method of image fusion based on bilateral filter is used in this experiment. The detail layer is extracted from a VIS image and fused

with the corresponding IR image by a strategy of average fusion.

The Qualitative results of edge map registration and image registration are shown in Fig. 5, where the edge maps of the IR images are blue and the edge maps of the VIS images are red. The registration results of CPD are not good on the image pair of *Block*. RGF is failure on the image pair of *Stairs*. RPM-L2E cannot achieve correct registration on the image pair of *Stairs* and *Building*. Meanwhile, all of the IR images are correctly registered to the corresponding VIS images by using our method, and the registration results of our method are better than those of the other methods on the whole tested image pairs, especially on the image pairs of *Stairs*, *Building* and *Block*. This is because the transformation models, which are estimated from point feature by using the strategy with global to local transformation, have strong generalization ability and are able to achieve image registration with global optima. Hence, the qualitative comparison demonstrates that our method is able to produce more accurate registration results consistently compared with the state-of-the-art methods.

Next, we give the quantitative comparisons of CPD, RGF, RPM-L2E, our method with the affine model and our method on the dataset. For each IR and VIS image pair, a set of point correspondence is constructed manually as ground truth which can be used to calculate the recall of registration results.

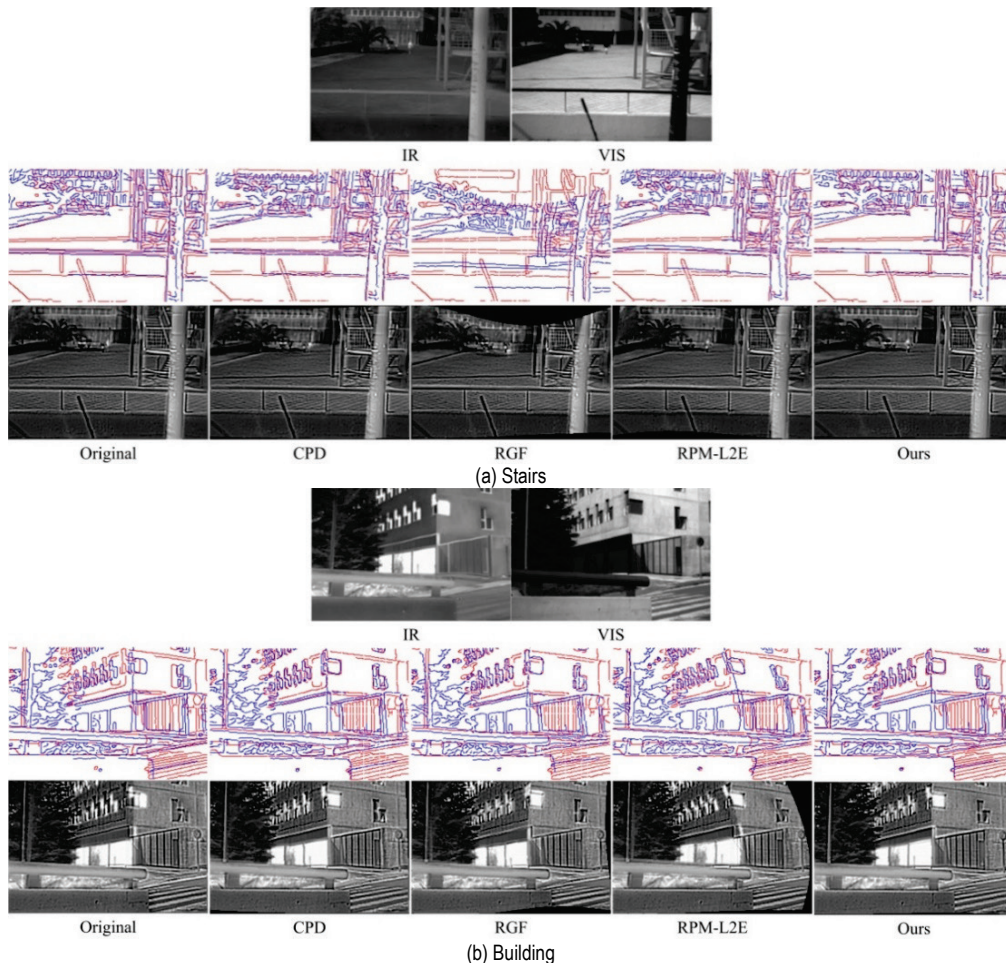


Figure 5 Qualitative comparisons of CPD, RGF, RPM-L2E and our method on the various image pairs, where the edge maps of the IR images are blue and the edge maps of the VIS images are red.

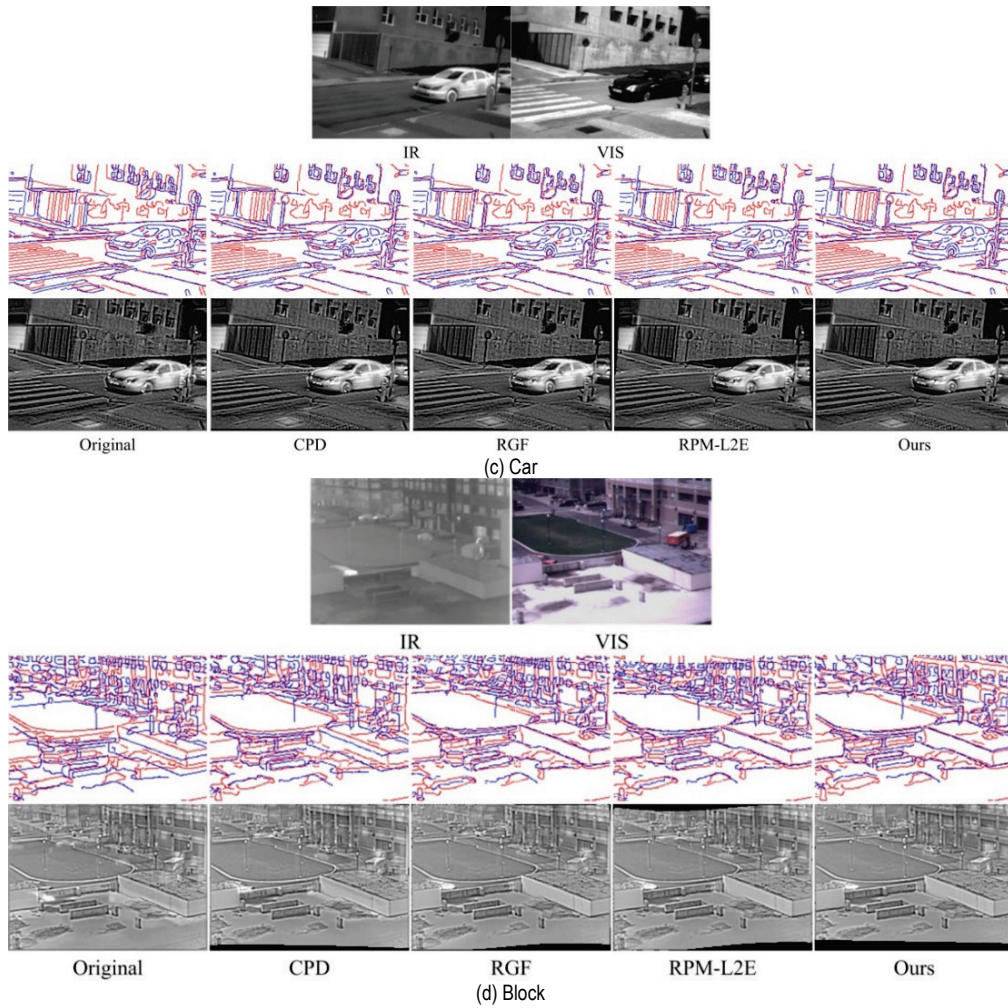


Figure 6 Qualitative comparisons of CPD, RGF, RPM-L2E and our method on the various image pairs, where the edge maps of the IR images are blue and the edge maps of the VIS images are red. (Continued from page 180)

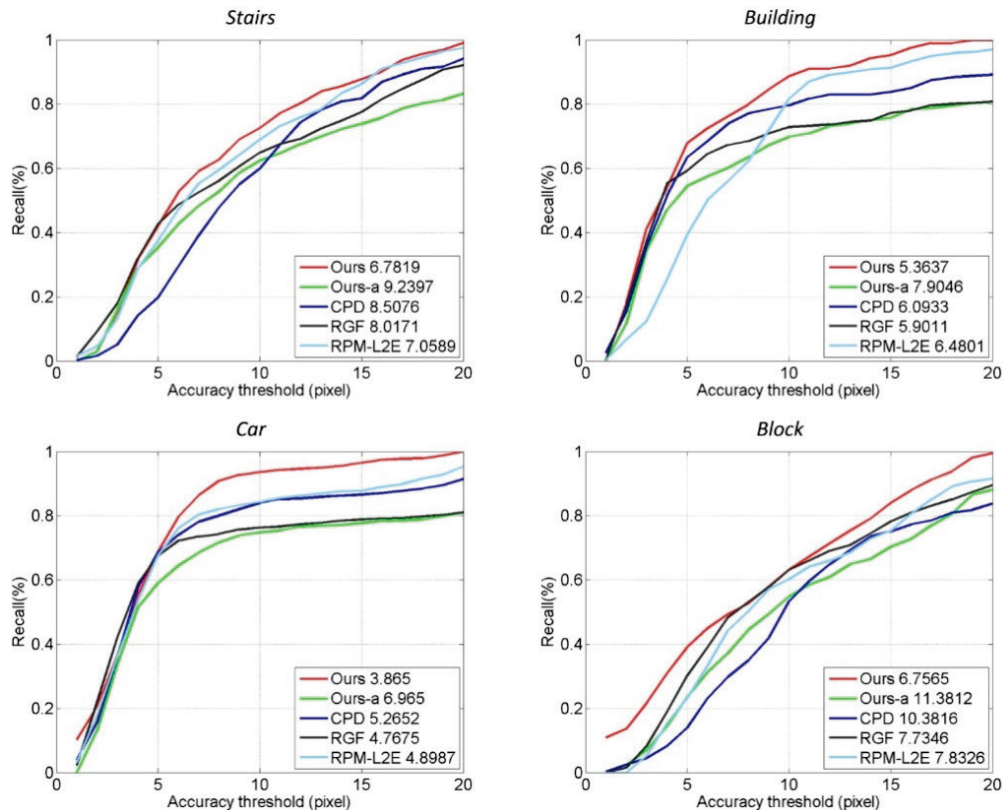


Figure 7 Quantitative comparisons of CPD, RGF, RPM-L2E, our method with the affine model (Ours-a) and our method on the various image pairs. The numbers in the legend are the average matching errors.

Fig. 6 reports the quantitative comparisons of the five methods on the four image pairs. As we can see the quantitative results are in good agreement with the qualitative results. The average matching errors of CPD, RGF, RPM-L2E and our method are about 7.56, 6.61, 6.57 and 5.69 pixels, respectively, while the average matching error of our method with the affine model is about 8.87 pixels. It proves the superiority of non-rigid transformation model for IR and VIS image registration. Meanwhile, our method has the least matching error and its recall curves are mainly above the curves of the other methods on the whole image pairs. This demonstrates that the proposed strategy with global to local transformation is able to improve the accuracy of IR and VIS image registration.

4 CONCLUSION

In this paper, a method with global to local transformation is proposed for non-rigid point registration and non-rigid registration of IR and VIS image. The global transformation model is introduced to describe the regular pattern of non-rigid deformation and achieve coarse registration in global scale. With the result of coarse registration as initial value, final point set registration is then completed by the fine registration with the local transformation model. To achieve non-rigid registration of IR and VIS images, the optimal transformation models are estimated from edge points by using the strategy with global to local transformation. The experiment on point sets and real images shows that the proposed method has good performance on non-rigid registration. The average matching error of the proposed method is decreased by at least 13.4% compared with the state-of-the-art methods.

Acknowledgments

The authors gratefully acknowledge the financial supports from the National Natural Science Foundation of China (No. 61901157), and the Science Foundation of Yunnan Electric Power Research Institute (Group) Co., Ltd. in China (No. YNKJXM20180244).

5 REFERENCES

- [1] Ma, J., Chen, C., Li, C., et al. (2016). Infrared and visible image fusion via gradient transfer and total variation minimization. *Information Fusion*, 31, 100-109. <https://doi.org/10.1016/j.inffus.2016.02.001>
- [2] Gómez, O., Ibáñez, O., Valsecchi, A., et al. (2018). 3D-2D silhouette-based image registration for comparative radiography-based forensic identification. *Pattern Recognition*, 83, 469-480. <https://doi.org/10.1016/j.patcog.2018.06.011>
- [3] Palanivel, D. A., Natarajan, S., & Gopalakrishnan, S. (2019). Multifractals based multimodal 3D image registration. *Biomedical Signal Processing and Control*, 47, 126-136. 2019. <https://doi.org/10.1016/j.bspc.2018.08.015>
- [4] Keane, T. P., Rhody, H. E., Saber, E., et al. (2012). Practical image registration concerns overcome by the weighted and filtered mutual information metric. *Journal of Electronic Imaging*, 21(2), 023029. <https://doi.org/10.1117/1.JEI.21.2.023029>
- [5] Zimmer, V. A., Gonzales, B., Miguel, A., et al. (2019). Multimodal image registration using Laplacian commutators. *Information Fusion*, 49, 130-145. <https://doi.org/10.1016/j.inffus.2018.09.009>
- [6] Budge, S. E., Badamkar, N. S., & Xie, X. (2014). Automatic registration of fused lidar/digital imagery (texel images) for three-dimensional image creation. *Optical Engineering*, 54(3), 031105. <https://doi.org/10.1117/1.OE.54.3.031105>
- [7] Tatano, R., Berkels, B., & Deserno, T. M. (2017). Mesh-to-raster region-of-interest-based nonrigid registration of multimodal images. *Journal of Medical Imaging*, 4(4), 044002. <https://doi.org/10.1117/1.JMI.4.4.044002>
- [8] Makrogiannis, S. and Bourbakis, N. G. (2010). Efficient registration of multitemporal and multisensor aerial images based on alignment of nonparametric edge features. *Journal of Electronic Imaging*, 19(1), 013002. <https://doi.org/10.1117/1.3293436>
- [9] Wei, Z., Tao, F., & Jun, W. (2013). Adaptive robust image registration approach based on adequately sampling polar transform and weighted angular projection function. *Journal of Electronic Imaging*, 22(4), 043030. <https://doi.org/10.1117/1.JEI.22.4.043030>
- [10] Müller, J., Müller, J., Chen, F., et al. (2018). Registration and fusion of thermographic and visual-light images in neurosurgery. *IEEE Transactions on Biomedical Circuits and Systems*, 12(6), 1313-1321. <https://doi.org/10.1109/TBCAS.2018.2856407>
- [11] Huang, H., Zhou, Z., & Cao, J. (2018). Non-rigid point set registration by high dimensional representation. *IET Image Processing*, 12(10), 1746-1752. <https://doi.org/10.1049/iet-ipr.2017.1363>
- [12] Stearns, C. C. & Kannappan, K. (1995). Method for 2-D affine transformation of images: U. S. Patent 5, 475, 803.
- [13] Bookstein, F. L. (1989). Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 567-585. <https://doi.org/10.1109/34.24792>
- [14] Xie, Z. & Farin, G. E. (2004). Image registration using hierarchical B-splines. *IEEE Transactions on Visualization and Computer Graphics*, 10(1), 85-94. <https://doi.org/10.1109/TVCG.2004.1260760>
- [15] Ma, J., Zhao, J., Ma, Y., et al. (2014). Non-rigid visible and infrared face registration via regularized Gaussian fields criterion. *Pattern Recognition*, 48, 772-784. <https://doi.org/10.1016/j.patcog.2014.09.005>
- [16] Liu, G., Liu, Z., Liu, S., et al. (2018). Registration of infrared and visible light image based on visual saliency and scale invariant feature transform. *EURASIP Journal on Image & Video Processing*, 2019, 45-57. <https://doi.org/10.1186/s13640-018-0283-9>
- [17] Lv, G. (2018). A novel correspondence selection technique affine rigid image registration. *IEEE Access*, 6, 32023-32034. <https://doi.org/10.1109/ACCESS.2018.2847399>
- [18] Raza, S., Sanchez, V., Prince, G., et al. (2015). Registration of thermal and visible light images of diseased plants using silhouette extraction in the wavelet domain. *Pattern Recognition*, 48(7), 2119-2128. <https://doi.org/10.1016/j.patcog.2015.01.027>
- [19] Ahmad, S. and Khan, M. F. (2018). Multimodal non-rigid image registration based on elastodynamics. *Visual Computer*, 34(1), 21-27. <https://doi.org/10.1007/s00371-016-1307-z>
- [20] Yang, C., Liu, Y., Jiang, X., et al. (2018). Non-rigid point set registration via adaptive weighted objective function. *IEEE Access*, 6, 75957-75960. <https://doi.org/10.1109/ACCESS.2018.2883689>
- [21] Chui, H. & Rangarajan, A. (2003). A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3), 114-141. [https://doi.org/10.1016/S1077-3142\(03\)00009-2](https://doi.org/10.1016/S1077-3142(03)00009-2)
- [22] Bong, A., Yang, B., Zhao, D., et al. (2013). Research of medical image non-rigid registration based on TPS-semisurf algorithm. *Advanced Materials Research*, 791-793, 2112-

2117.
<https://doi.org/10.4028/www.scientific.net/AMR.791-793.2112>
- [23] Yang, C., Liu, Y., Jiang, X., et al. (2018). Non-rigid point set registration via adaptive weighted objective function. *IEEE Access*, 6, 75947-75960.
<https://doi.org/10.1109/ACCESS.2018.2883689>
- [24] Gu, S., Meng, X., Sciurba, F. C., et al. (2014). Bidirectional elastic image registration using B-spline affine transformation. *Computerized Medical Imaging and Graphics*, 38, 306-314.
<https://doi.org/10.1016/j.compmedimag.2014.01.002>
- [25] Jiang, P. & Shackelford, J. A. (2015). B-spline registration of neuroimaging modalities with map-reduce framework. *Brain Informatics and Health*, 9250, 285-294.
https://doi.org/10.1007/978-3-319-23344-4_28
- [26] Sun, W., Niessen, W. J., & Klein, S. (2017). Randomly perturbed B-splines for nonrigid image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(7), 1401-1413.
<https://doi.org/10.1109/TPAMI.2016.2598344>
- [27] Ma, J., Qiu, W., Zhao, J., et al. (2015). Robust L₂E estimation of transformation for non-rigid registration. *IEEE Transactions on Signal Processing*, 63(5), 1115-1127.
<https://doi.org/10.1109/TSP.2014.2388434>
- [28] Ma, J., Zhao, J., & Yuille, A. L. (2016). Non-rigid point set registration by preserving global and local structures. *IEEE Transactions on Image Processing*, 25(1), 53-64.
<https://doi.org/10.1109/TIP.2015.2467217>
- [29] Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4), 509-522. <https://doi.org/10.1109/34.993558>
- [30] Liu, X., Ai, Y., Zhang, J., et al. (2018). A novel affine and contrast invariant descriptor for infrared and visible image registration. *Remote Sensing*, 10, 658.
<https://doi.org/10.3390/rs10040658>

Contact information:**Hao PAN**

(Corresponding author)
 Electrical Research Institute of Yunnan Electric Power Research Institute
 (Group) Co., Ltd. Kunming, 650051, China
 E-mail: panhaohu@yeah.net

Yi MA

Electrical Research Institute of Yunnan Electric Power Research Institute
 (Group) Co., Ltd. Kunming, 650051, China

Fangrong ZHOU

Electrical Research Institute of Yunnan Electric Power Research Institute
 (Group) Co., Ltd. Kunming, 650051, China

Yan GU

North Night Vision Technology Corp., Ltd. Nanjing, 211102, China

Yutang MA

Electrical Research Institute of Yunnan Electric Power Research Institute
 (Group) Co., Ltd. Kunming, 650051, China

Chaobo MIN

North Night Vision Technology Corp., Ltd. Nanjing, 211102, China