



# Non-rigid point set registration using dual-feature finite mixture model and global-local structural preservation



Su Zhang<sup>a,b,c</sup>, Kun Yang<sup>a,b,\*</sup>, Yang Yang<sup>a,b,c,\*</sup>, Yi Luo<sup>a,b</sup>, Ziquan Wei<sup>a,b,c</sup>

<sup>a</sup> School of Information Science and Technology, Yunnan Normal University, Kunming 650092, China

<sup>b</sup> The Engineering Research Center of GIS Technology in Western China of Ministry of Education, 650500, China

<sup>c</sup> The Laboratory of Pattern Recognition and Artificial Intelligence, Yunnan Normal University, Kunming, Yunnan 650092, China

## ARTICLE INFO

### Article history:

Received 13 January 2017

Revised 15 January 2018

Accepted 4 March 2018

Available online 17 March 2018

### Keywords:

Registration

Dual-feature

Finite mixture model

Regularization

Structural preservation

## ABSTRACT

We present a dual-feature based point set registration method with global-local structural preservation. A finite mixture model which is able to deal with two features is first constructed. The mixture structure descriptor (MSD) is then obtained by smoothly combining the local structure descriptor (LSD) with the original coordinates through an annealing scheme. Substituting the MSD into the constructed model a fuzzy corresponding matrix is acquired. Next, the energy function which contains three main terms is formulated in the reproducing kernel Hilbert space. The first term is based on the  $L_2$  estimation ( $L_2E$ ) criterion, and the other two terms play a complementary role to improve the robustness and accuracy for transformation estimation at both global and local scales. The performances of our method in synthetic data, sequence data and real data against nine state-of-the-art methods are tested where our method shows favorable performance when compared with other methods.

© 2018 Elsevier Ltd. All rights reserved.

## 1. Introduction

Non-rigid point set registration [1] is the process of aligning one set of points (the source point set) to another (the target point set) which is deformed from the former. It is not only important for many computer vision tasks such as stereo matching, content-based image retrieval, image registration [2–5], object detection [6], but also widely applied in other fields like mobile robotics [7], medical image analysis [8,9], natural disaster damage assessment [10], environment changes assessment [11] and remote sensing processing [12,13]. According to the methodological differences of current popular non-rigid point set registration methods, there are two major types as follows: (i) iterative or non-iterative methods, and (ii) learning or non-learning based methods. We introduce and discuss the current methods which belongs to type (i), since developing an iterative method is the focus of this paper.

Iterative methods repeat two major steps which are (i) correspondence estimation and (ii) transformation estimation until convergence or reaching the maximum iteration number. The key idea of iterative methods is to gradually adjust the initial geometrical structure and the location of the source point set, thereby making it easier to estimate correspondence. The iterated closest point

(ICP) [14] algorithm is one of the best known algorithms because of its simplicity and low computational complexity. The thin-plate spline robust point matching (TPS-RPM) algorithm [15] uses several techniques such as soft assign [16], deterministic annealing [17], TPS [18] transformation and Tikhonov regularization framework [19]. GMMREG [20] introduces a Gaussian mixture model approach for both rigid and non-rigid registration by minimizing the closed-form expression of  $L_2$  distance between two Gaussian mixture models. Nicholas et al. [21] introduced a labeled point set registration algorithm based on the Jensen–Havrda–Charvat–Tsallis (JHCT) divergence measure. Myronenko and Song [22] proposed the CPD algorithm, which was a maximum likelihood estimation of Gaussian mixture model with a motion coherence constraint [23] for preserving the global structure of the point sets. Moreover, the fast Gaussian transform [24] (FGT) and low-rank matrix approximation [25] were adopted to provide a fast implementation. RPM-L2E [26] employs the shape context [27] as the feature descriptor to estimate the correspondence, then the transformation is refined by using the  $L_2$  estimation ( $L_2E$ ) [28] criterion, where a sparse approximation to further speed-up the computation is applicable. Recently, GLMDTPS [29] presents a multi-feature based correspondence estimation which is named as global and local mixture distance (GLMD). MoAGREG [30] introduces asymmetric Gaussian model [31] to represent the density of the given points.

\* Corresponding authors.

E-mail address: [yyang\\_ynu@163.com](mailto:yyang_ynu@163.com) (Y. Yang).

**Table 1**

Methodological differences on correspondence estimation and outlier dealing between our method and other three methods. GPD: Gaussian probability density; SC: shape context [27]; GLMD: global and local mixture distance; MSD: mixture structure descriptor; DFMMMPD: dual-feature finite mixture model probability density; UD: uniform distribution;  $L_2E$ : the  $L_2E$  estimator [28].

Method	Correspondence estimation		Outlier dealing
	Using feature	Correspondence	
CPD	GPD	Fuzzy	UD
RPM-L2E	SC	Binary	$L_2E$
GLMDTPS	GLMD	Binary	✗
Ours	DFMMMPD + MSD	Fuzzy	UD+ $L_2E$

Of all the aforementioned iterative methods, there exists two problems : (i) ICP, TPSRPM, GMMREG, CPD, RPM-L2E and MoA-GREG estimate the correspondence based on single feature (e.g., the point-wise Euclidean distance or the shape context descriptor); (ii) all the methods only use single regularization on their radial basis functions for preserving global structural consistency in transformation estimation, therefore it is difficult to give reliable results on different registration patterns.

Motivated by the above two problems, a dual-feature based point set registration method with global-local correspondence and transformation estimations is presented. (i) The dual-feature finite mixture model, which works together with the point-wise Euclidean distance and the local structure descriptor (LSD) discrepancies for obtaining the reliable correspondence estimation, is first constructed. (ii) By extending the LSD to be the local structure preservation (LSP) term, the respective advantages of the MCT based regularization and the LSP complement each other, forming the global-local structural preserved  $L_2E$  energy function.

The rest of the paper is organized as follows: Section 2 discusses the methodological differences of our method with the most related works; Section 3 details our method; in Section 4, the performances of our method in abundant experiments are tested, and the results of our method versus seven state-of-the-art are reported; finally, the conclusion is provided in Section 5.

## 2. Related works

There are three existing methods: CPD [22], RPM-L2E [26] and GLMDTPS [29] that are most relevant to our method. We discuss the methodological differences between our method and other three methods from five aspects: (i) correspondence estimation; (ii) transformation estimation; (iii) outlier dealing; (iv) covariances; (v) regularization softening.

### 2.1. Correspondence estimation

CPD considers point set registration as a probability density estimation problem, and it takes the source and target point set as Gaussian components and data, respectively. Since CPD uses expectation maximization (EM) algorithm to optimize the energy function, a posterior probability matrix is constructed based on Bayes' theorem, which acts as a corresponding matrix. RPM-L2E employs the shape context to estimate correspondence. Based on the shape distance matrix, the Hungarian algorithm is used to solve the linear assign problem, and a binary matrix is therefore obtained as the correspondence. A similar one-to-one corresponding matrix is obtained in GLMDTPS by optimizing the mixture distance matrix using the Jonker-Volgenant algorithm [32]. For CPD and RPM-L2E, the correspondence estimation is determined by the single feature. In this paper, the global and local features are incorporated into the a dual-feature finite mixture model (DFMM), by which a fuzzy corresponding matrix is yielded (Table 1).

**Table 2**

Methodological differences on transformation updating between our method and other three methods. GRBF: Gaussian radial basis function; TPS: thin-plate spline; G: global structure preservation; L: local structure preservation.

Method	Transformation updating		
	Mapping	Preservation	Softening
CPD	GRBF	G	✗
RPM-L2E	GRBF	G	✗
GLMDTPS	TPS	G	✓
Ours	GRBF	G+L	✓

### 2.2. Transformation estimation

All of the four methods model the non-rigid transformation in the reproducing kernel Hilbert space (RKHS)  $\mathcal{H}$ . Typically, the Gaussian kernel and the TPS kernel are chosen for defining a  $\mathcal{H}$  uniquely, whereby the transformation models are determined as the Gaussian radial basis functions (GRBF) and thin-plate spline (TPS), respectively. CPD and RPM-L2E use the GRBF, while GLMDTPS uses the TPS. All these methods regularize their energy function by the same criterion, i.e., minimizing the spatial deviations between the corresponding points while forcing all the points from the source set to move coherently at global scale. In this paper, a new regularization term named local structure preservation (LSP) is formulated. The LSP can preserve neighborhood structures of the points during the transformation. Intuitively, for two corresponding point, their local structures defined by the K nearest neighbors are imposed to align with each other (Table 2).

### 2.3. Outlier dealing

CPD considers the outlier rejection as an unsupervised clustering problem, and RPM-L2E uses the robust  $L_2E$  estimator for outlier dealing. In our method, outliers are evaluated by the uniform distribution contained in the DFMM as well as the  $L_2E$  estimator in transformation estimation.

### 2.4. Covariances

CPD uses EM algorithm, the covariances are updated in M-step by taking the derivative of the objective function. RPM-L2E adopts Quasi-Newton method to update the transformation, the covariances are initialized as the same values, which are then annealed for fixed times during each iteration. Two covariances are included in our method. The former are the searching ranges of the DFMM, which are progressively decreased in each iteration. The latter belong to the  $L_2E$  based energy function, which are initialized as the same values in each iteration, and then decrease in each sub-iteration.

### 2.5. Regularization softening

GLMDTPS uses the softening technique in their TPS based energy function. In our method, the global and local preservation terms are softened by a well defined function. It remains stable in most iterations, then sharply decreases to zero in the last several iterations. And for CPD and RPM-L2E, they use the constant regularizations.

## 3. Method

In this section,<sup>1</sup> the main process of our method is first introduced, whereafter the major contributions in the correspondence

<sup>1</sup> The abbreviations throughout Section Method are listed here. CE: correspondence estimation; DFMM: dual-feature finite mixture model; GLSP: global-local

and transformation estimations are detailed. Finally, the pseudo-code and the analysis of computational complexity are provided.

### 3.1. Main process

Given two point sets, the source point set  $\mathbf{Y} = \{\mathbf{y}_j | \mathbf{y}_j \in \mathbb{R}^D, j = 1, \dots, M\}$  and the target point set  $\mathbf{X} = \{\mathbf{x}_i | \mathbf{x}_i \in \mathbb{R}^D, i = 1, \dots, N\}$ , where  $M \leq N$  and  $\mathbf{X}$  is typically perturbed by deformation, noise, outlier and rotation. The goal is to recover the unknown non-rigid transformation  $\mathcal{T}^* = \{\mathcal{T}^*(\mathbf{y}_1), \mathcal{T}^*(\mathbf{y}_2), \dots, \mathcal{T}^*(\mathbf{y}_M)\}^T$  iteratively which registers  $\mathbf{Y}$  to  $\mathbf{X}$ , hopefully resulting in the maximum point-wise overlap within inliers. The main process consists of an alternating two-step process: (i) estimating correspondence between  $\mathbf{Y}$  and  $\mathbf{X}$  as an initialization, thus the putative target point set  $\hat{\mathbf{X}} = \{\hat{\mathbf{x}}_i | \hat{\mathbf{x}}_i \in \mathbb{R}^D, i = 1, \dots, M\}$  in the current iteration is obtained, where  $\mathbf{Y}$  and  $\hat{\mathbf{X}}$  comprise a set of correspondences  $\{\mathbf{y}_i, \hat{\mathbf{x}}_i\}_{i=1}^M$ ; (ii) the point-wise registration between  $\mathbf{Y}$  and  $\hat{\mathbf{X}}$  is yielded by finding a non-rigid transformation  $\mathcal{T} : \hat{\mathbf{X}} = \mathcal{T}(\mathbf{Y})$ , where  $\mathbf{Y}$  is updated:  $\mathbf{Y} \leftarrow \mathcal{T}(\mathbf{Y})$  and then substituted into step (i) in the next iteration.

#### 3.1.1. Correspondence Estimation (CE)

The CE of  $\mathbf{Y}$  and  $\mathbf{X}$  is considered as a probability density estimation problem. Instead of using Gaussian mixture model which exploits only the global point-wise Euclidean distance [20,22], a dual-feature finite mixture model (DFMM) is developed, by which both the global and local features can be utilized well for a reliable CE.

Let  $\mathbf{y}_j$  be the centroid of the  $j$ th dual-feature Gaussian component,  $\mathbf{x}_i$  the  $i$ th data. For all components (i.e.,  $j = 1, 2, \dots, M$ ), equal isotropic covariances  $\sigma^2$  and prior probabilities  $C_j = 1/M$  are applied. The probability density function of the DFMM is written as:

$$p(\mathbf{x}_i) = (1 - \omega) \sum_{j=1}^M C_j p(\mathbf{x}_i | \mathbf{y}_j, \sigma^2 \mathbf{I}) + \omega p(\mathbf{x}_i | M + 1), \quad (1)$$

where  $\mathbf{I}_{D \times D}$  is an identity matrix.  $p(\mathbf{x}_i | \mathbf{y}_j, \sigma^2 \mathbf{I})$  implies that what proportion of  $\mathbf{x}_i$  is captured by  $\mathbf{y}_j$  within the searching range  $\sigma^2$ . An uniform distribution  $p(\mathbf{x}_i | M + 1) = 1/N$  is employed for outlier modeling with a weighting parameter  $\omega$ . Thus, the correspondence between  $\mathbf{y}_j$  and  $\mathbf{x}_i$  can be easily inferred through the posterior probability [22] following the Bayes' theorem as:

$$P(\mathbf{y}_j | \mathbf{x}_i) = \frac{C_j p(\mathbf{x}_i | \mathbf{y}_j, \sigma^2 \mathbf{I})}{\sum_{k=1}^M C_k p(\mathbf{x}_i | \mathbf{y}_k, \sigma^2 \mathbf{I}) + \frac{\omega}{(1-\omega)N}}. \quad (2)$$

These posterior probabilities are constrained such that  $\sum_{j=1}^M P(\mathbf{y}_j | \mathbf{x}_i) = 1$ . Based on the assumption that each source point corresponds to a weighted sum of the target points, the putative target point set is obtained by  $\hat{\mathbf{X}} = \mathbf{P}\mathbf{X}$ , where  $\mathbf{P}_{M \times N}$  is the posterior probability matrix whose element in the  $j$ th row,  $i$ th column is  $P(\mathbf{y}_j | \mathbf{x}_i)$ . Moreover, the covariance parameter  $\sigma^2$  is directly controlled through deterministic annealing, whereby the values of the posterior matrix  $\mathbf{P}$  are nearly uniform at early stage of iterations, then close to binary as the iteration number increases.

#### 3.1.2. Transformation Estimation (TE)

Subsequent to CE, we suppose that the noise on the inliers is Gaussian with zero mean and uniform standard deviation  $\rho$ , or to

be more specifically, an inlier point correspondence  $(\hat{\mathbf{x}}_i, \mathbf{y}_i)$  satisfies  $\hat{\mathbf{x}}_i - \mathcal{T}(\mathbf{y}_i) \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$ . The data  $\{\hat{\mathbf{x}}_i - \mathcal{T}(\mathbf{y}_i)\}_{i=1}^M$  can be taken as a sample set from an outlier-contaminated multivariate normal density  $\mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$ . Hence, the energy function is in the form

$$E_{L_2 E}(\mathcal{T}, \rho^2) = \frac{1}{2^d (\pi \sigma_2)^{\frac{D}{2}}} - \frac{2}{M} \sum_{m=1}^M \Phi(\hat{\mathbf{x}}_i - \mathcal{T}(\mathbf{y}_i) | \mathbf{0}, \rho^2 \mathbf{I}) \quad (3)$$

based on  $L_2$ -minimizing estimate ( $L_2 E$ ) criterion [26], where  $\Phi$  is a multi-dimensional Gaussian function and  $\mathbf{I}_{D \times D}$  is an identity matrix. Minimizing  $E_{L_2 E}$  is equivalent to registering  $\mathbf{y}_i$  to  $\hat{\mathbf{x}}_i$  so that the pairwise spatial deviation of the correspondence set  $\{\hat{\mathbf{x}}_i, \mathbf{y}_i\}_{i=1}^M$  is minimized.

The key constraint of a rigid transformation is that any point-wise distances are preserved. However, once non-rigidity is allowed, the distance between a pair of points will, in general, not be preserved, which leads to ill-posed problems since the solution of  $\mathcal{T}$  might not be unique. In order to deal with the problem, the Tikhonov regularization framework [19] is employed, and the energy function is then written in the form:

$$E(\mathcal{T}, \rho^2) = E_{L_2 E}(\mathcal{T}, \rho^2) + \mathcal{R}(\mathcal{T}), \quad (4)$$

where term  $\mathcal{R}(\mathcal{T})$  is the regularization on  $\mathcal{T}$ . In our method, the unknown non-rigid transformation  $\mathcal{T}$  is formulated in reproducing kernel Hilbert space (RKHS) [33], and the regularization  $\mathcal{R}(\mathcal{T})$  is defined at both the global and local scales. Efficient gradient-based numerical optimization techniques, e.g., the quasi-Newton method and the nonlinear conjugate gradient method are applicable for solving the optimization problem on Eq. (4) updating  $\mathbf{Y} \leftarrow \mathcal{T}(\mathbf{Y})$ . Next, we return to the CE step and continue the registration process until the maximum iteration number is reached.

### 3.2. Improvement of CE: DFMM construction and feature complementation

For a Gaussian component  $p(\mathbf{x}_i | \mathbf{y}_j, \sigma^2 \mathbf{I}) = \frac{1}{(2\pi\sigma^2)^{D/2}} \exp(-\|\mathbf{x}_i - \mathbf{y}_j\|^2 / 2\sigma^2)$ , it only considers the Euclidean distance  $\|\mathbf{x}_i - \mathbf{y}_j\|^2$ , which results in insufficient robustness in many registration patterns. For instance, if  $\|\mathbf{x}_a - \mathbf{y}_j\|^2 = \|\mathbf{x}_b - \mathbf{y}_j\|^2$ , equal posterior probabilities are obtained, although the local geometrical structure around  $\mathbf{x}_a$  and  $\mathbf{x}_b$  are different. To cope with this problem, the dual-feature finite mixture model is developed, to which the global and local features with an appropriate technique for combination and complementation are integrated.

#### 3.2.1. Construction of the dual-feature mixture model (DFMM)

The density of a DFMM component is defined as:

$$p(\mathbf{x}_i | \mathbf{y}_j, \sigma^2 \mathbf{I}) = \frac{1}{(2\pi\sigma^2)^{\frac{D}{2}}} \exp\left(-\frac{\|\Delta_{ij}^G + \alpha \Delta_{ij}^L\|^2}{2\sigma^2}\right), \quad (5)$$

where  $\Delta_{ij}^G$  and  $\Delta_{ij}^L$  denote the global and local geometrical discrepancies between  $\mathbf{y}_j$  and  $\mathbf{x}_i$ , respectively, with the parameter  $\alpha$  denoting the trade-off.

The underlying assumption of density function Eq. (5) is the decomposition of the process for human to recognize and categorize objects. Supposing such process is based on the linear combination among features such as Euclidean distance and density, etc. The priority of certain feature may change during the process. For instance, one can easily categorize different letters according to the feature of shape at the very beginning, after which the accuracy can be further optimized by involving other features in. In our DFMM, the correspondence is estimated according to the discrepancies of the point-wise Euclidean distance and the local structure descriptor (LSD), and the interchange of these two features is controlled by the parameter  $\alpha$ .

structural preservation; KNN: K nearest neighbors; L2E:  $L_2$  estimation criterion; LSD: local structure descriptor; LSP: local structure preservation; MLE: maximum log-likelihood estimator; MSD: mixture structure descriptor; ODR: outlier to data ratio; RKHS: reproducing Hilbert kernel space; TE: transformation estimation.

### 3.2.2. Feature combination and complementation

To complement the dubious point-wise Euclidean distance in CE, the local structure descriptor (LSD) is defined as:

$$L_{\mathbf{x}_i} = \sum_{k=1}^K w_{ik} (\mathbf{x}_{ik} - \mathbf{x}_i) \quad (6)$$

where  $(\mathbf{x}_{ik} - \mathbf{x}_i)$  is actually a vector from  $\mathbf{x}_i$  to its  $k$ th nearest neighbors  $\mathbf{x}_{ik}$ . Considering the vector may be too high a value in some cases, and in general, the neighbors of  $\mathbf{x}_i$  are distributed normally, as mentioned earlier, the parameter  $w_{ik} = \exp(-\|\overrightarrow{\mathbf{x}_i \mathbf{x}_{ik}}\|^2/v_i^2)$  is therefore used to weight the contribution of the  $k$ th local vector to the LSD, where  $v_i$  is the variance of the vector norms  $\{\|\mathbf{x}_{ik} - \mathbf{x}_i\|\}_{k=1}^K$ .

Next, the Euclidean coordinates of  $\mathbf{y}_j$  and  $\mathbf{x}_i$  are replaced by  $\mathcal{M}\mathbf{y}_j = \mathbf{y}_j + \alpha L_{\mathbf{y}_j}$  and  $\mathcal{M}\mathbf{x}_i = \mathbf{x}_i + \alpha L_{\mathbf{x}_i}$ , where  $\mathcal{M}$  denotes the mixture structure descriptor (MSD), and  $\alpha$  is defined as  $\alpha = \exp(-t/r)$ , where  $t$  is the current iteration number, and  $r$  is a constant. The use of the MSD is analogous to projecting the point from the Euclidean space onto its MSD space whose element entails the global-local structural information.

Substituting the MSDs into Eq. (2), the complete posterior probability function is obtained as:

$$p_{ij} = \frac{\exp\left(-\frac{\|\Delta_{ij}^G + \alpha \Delta_{ij}^L\|^2}{2\sigma^2}\right)}{\sum_{k=1}^M \exp\left(-\frac{\|\Delta_{ik}^G + \alpha \Delta_{ik}^L\|^2}{2\sigma^2}\right) + \frac{(2\pi\sigma^2)^{\frac{D}{2}}\omega M}{(1-\omega)N}}, \quad (7)$$

where  $\Delta_{ij}^G = \mathbf{x}_i - \mathbf{y}_j$ , and  $\Delta_{ij}^L = L_{\mathbf{x}_i} - L_{\mathbf{y}_j}$ . A deterministic annealing process is performed by gradually lowering the covariance  $\sigma^2$  by  $\sigma^2 \leftarrow \epsilon \sigma^2$  as the iteration number increases, where  $\epsilon$  is a constant.

The idea of adopting the annealing is based on the observations that when  $\sigma^2$  is large, the exponential term in Eq. (7) is close to zero, the resulting estimation of correspondence  $p_{ij}$  are then nearly uniform which indicates an one-to-many matching strategy (including those MSDs which are far away) with a fuzzy degree of confidence. When the magnitude of  $\sigma^2$  becomes small,  $p_{ij}$  is gradually close to binary, and the points are only allowed to match to the one whose MSD is the closest. These observations suggests that  $\sigma^2$  is basically performing like a searching range parameter. A direct control on  $\sigma^2$ , which results in a tunable fuzziness of the correspondence, is therefore considerable for enhancing the robustness of the CE in our method.

On the other hand, the definition of the gradually decreasing weight  $\alpha$  in each MSD also accommodates the annealing scheme. At the early stage of iterations, the point sets  $\mathbf{Y}$  and  $\mathbf{X}$  have the biggest difference, whereas the local relationship among neighboring points is stronger and more stable than the global one [34], the correspondence is therefore estimated through the LSD at this stage. By using large  $\alpha$ , the DFMM-based CE can elicit a good initial pose for the source point set  $\mathbf{Y}$ , this greatly facilitates the registration process at the very beginning. When iteration is large,  $\mathbf{Y}$  and  $\mathbf{X}$  are very similar, a straightforward estimation based on the Euclidean distance is desirable, this is achieved since the MSDs of the two point set are nearly converging to the Euclidean coordinates.

Overall, the DFMM allows us to achieve the complementation of the global and local features during the registration, and therefore gives accurate results. A demonstration of the registration processes using the point-wise Euclidean distance-based CE (in Euclidean space) and MSD-based CE (in MSD space), as well as the process for the MSDs to converge to the Euclidean coordinates are shown in Fig. 1.

### 3.3. Improvement of TE: $L_2E$ based energy function and global-local structural preservation

Once the putative target point set  $\hat{\mathbf{X}}$  is obtained using the MSD-based CE, our next concern is to find a non-rigid transformation  $\mathcal{T}: \hat{\mathbf{X}} = \mathcal{T}(\mathbf{Y})$ , which minimizes the point-wise spatial deviations between the source  $\mathbf{Y}$  and the putative target  $\hat{\mathbf{X}}$ . In our method,  $\mathcal{T}$  is parameterized using the Gaussian radial basis function, and solved by optimizing Eq. (4), the  $L_2E$  [28] criterion-based energy function, which is regularized by the global-local structural preservation term  $\mathcal{R}(\mathcal{T})$ .

#### 3.3.1. $L_2$ estimation ( $L_2E$ )

$L_2E$  is a robust estimator which minimizes the  $L_2$  distance between densities, and is particularly appropriate for analyzing massive datasets where outlier rejecting is impractical.

Suppose we have a density model  $p(z|\theta)$ , our goal is to minimize the estimate of  $L_2$  distance with respect to  $\theta$  as  $\theta^* = \arg \min_{\theta} \int [p(z|\theta) - p(z|\theta_0)]^2 dz$ , where the true parameter  $\theta_0$  is unknown. After omitting the constant  $\int p(z|\theta_0)^2 dz$ , the parameter  $\theta$  is estimated by minimizing the  $L_2E$  criterion as:

$$\theta_{L_2E}^* = \arg \min_{\theta} \left[ \int p(z|\theta)^2 dz - \frac{2}{m} \sum_{i=1}^m p(z|\theta) \right]. \quad (8)$$

To demonstrate the robustness of the  $L_2E$  estimator compared against the maximum log-likelihood estimator ( $MLE$ ), two series of experiments are conducted on a line-fitting problem as Fig. 2 shows. The inliers are generated from a linear regression model, i.e.,  $y = 0.5x + \epsilon$ , where the residual  $\epsilon \sim \mathcal{N}(0, 1)$ . The outliers are added (i) from another linear model, i.e.,  $y = 3.5x + \epsilon$ , or (ii) irregularly. The goal is to estimate the slope  $k$  of the inlier linear model using  $MLE$  and  $L_2E$ . For the two series, the number of the data ( $N = 200$ ) is fixed and the outlier to data ratio ( $ODR$ ) is varied as 10%, 20%, 40%, 60%, 100%. In this context, the optimal  $\hat{k}$  is obtained by solving  $k_{MLE} = \arg \max_k \sum_{i=1}^N \log \phi(y_i - kx_i | 0, 1)$  and  $k_{L_2E} = \arg \min_k [1/2\sqrt{\pi} - \frac{2}{N} \sum_{i=1}^N \phi(y_i - kx_i | 0, 1)]$ , respectively. In series (i), the  $L_2E$  always have the global optima at  $k = 0.5$ . And the local optima are at  $k = 3.5$ , which becomes deeper as the  $ODR$  increases. This is reasonable since the outliers are from the linear model with slope  $k = 3.5$ , while the two optima become approximately equal when  $ODR = 100\%$ . In series (ii), the optima of the  $L_2E$  are also near  $k = 0.5$  stably. However, the  $MLE$  in the two series are not resistant to outliers, since they deviate more under heavier contamination.

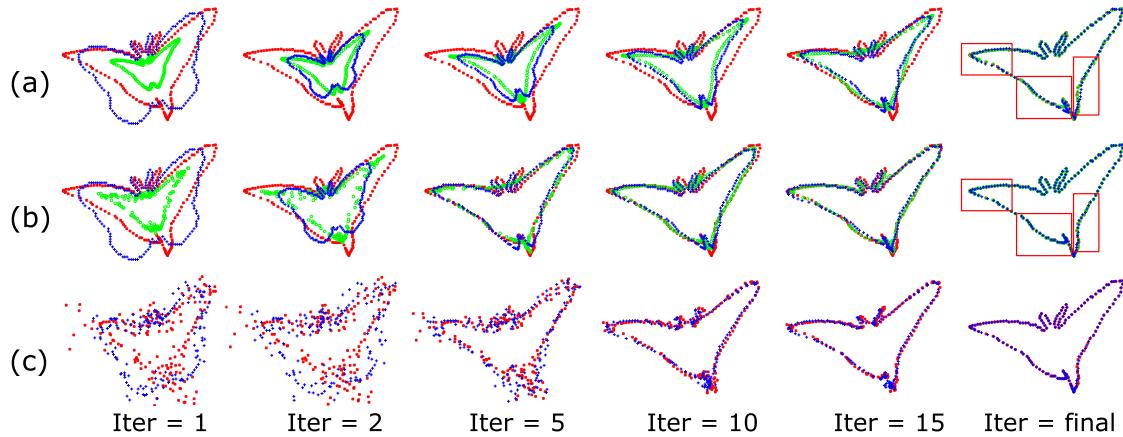
We now apply the  $L_2E$  in the point set registration problem and rewrite Eq. (4) following the  $L_2E$  criterion:

$$E(\mathcal{T}, \rho^2) = \frac{1}{2^d (\pi \rho)^{\frac{D}{2}}} - \frac{2}{M} \sum_{m=1}^M \frac{1}{(2\pi \rho^2)^{\frac{D}{2}}} \times \exp\left(-\frac{\|\hat{\mathbf{x}}_i - \mathcal{T}(\mathbf{y}_i)\|^2}{2\rho^2}\right) + \mathcal{R}(\mathcal{T}). \quad (9)$$

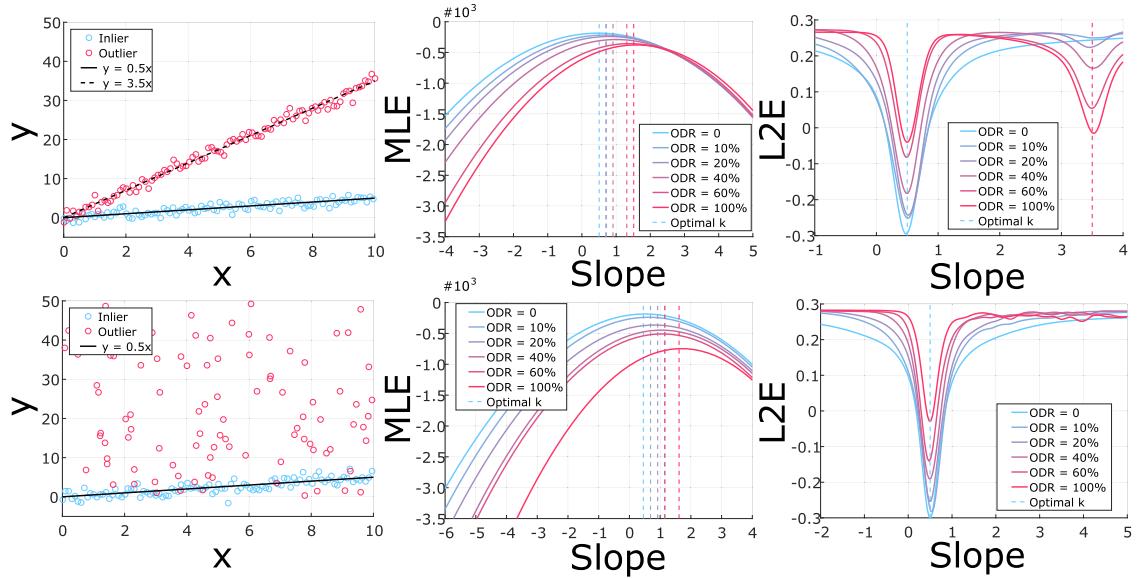
The transformation is defined as the initial position plus a displacement function  $\mathcal{V}$ :

$$\mathcal{T}(\mathbf{Y}, \mathcal{V}) = \mathbf{Y} + \mathcal{V}(\mathbf{Y}), \quad (10)$$

where  $\mathcal{V} = \{\mathcal{V}(\mathbf{y}_1), \mathcal{V}(\mathbf{y}_2), \dots, \mathcal{V}(\mathbf{y}_M)\}^T$ , and model the non-rigid transformation by requiring it to lie within a projected high-dimensional feature space, i.e., reproducing kernel Hilbert space (RKHS) [33]. Let  $\mathcal{K}: \mathbf{Y} \times \mathbf{Y} \mapsto \mathbb{R}^{D \times D}$  be a positive definite matrix-valued kernel, e.g., the Gaussian kernel, with an associated RKHS family of functions  $\mathcal{H}_{\mathcal{K}}$  and its corresponding norm  $\|\cdot\|_{\mathcal{H}}$ , then a useful property of RKHS is shown as follows:



**Fig. 1.** The registration examples of our method on the effect of LSD. Row (a): the registration examples using the Euclidean space-based CE. Row (b): the registration examples using the MSD space-based CE. Row (c): the visualization of the two point sets in the MSD space. In Row (b) and (c), Blue +: the source point set  $\mathbf{Y}$ . Red \*: the target point set  $\mathbf{X}$ . Green  $\circ$ : the estimated putative target point set  $\hat{\mathbf{X}}$ . The obvious registration errors are highlighted using the red rectangles. At the 1th iteration, we see that the initial  $\hat{\mathbf{X}}$  in row (b) is close to  $\mathbf{X}$ , while it is close to the regional centers of mass in row(a). Finally, the performance of row (b) outperforms. In row (c), the coordinates of  $\mathbf{X}$  and  $\mathbf{Y}$  in the MSD space converge to their Euclidean coordinates at later stage of iterations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 2.** The comparison between  $L_2E$  and  $MLE$  on the line-fitting problem. Two series of experiments are designed with each row showing one series. Left column: data samples, where the blue and red circles denote the inliers and outliers, respectively. The solid line and dash line represent the linear model for generating the inliers and outliers, respectively. Middle and right column: the fitting results by using the  $MLE$  and  $L_2E$ . The color gradient is used to show the different ODRs, where the light blue denotes  $ODR = 0$ , and the vermilion denotes  $ODR = 100\%$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Theorem 1.** (Riesz Representer Theorem) The minimization of the energy function Eq. (9) has an optimal solution, given by

$$\mathcal{V}(\mathbf{y}_i) = \sum_{j=1}^M \mathcal{K}(\mathbf{y}_i, \mathbf{y}_j) \mathbf{c}_j. \quad (11)$$

Inspired by the above theorem [35], the expression of the non-rigid transformation is obtain as:

$$\mathcal{T}(\mathbf{Y}, \mathbf{C}) = \mathbf{Y} + \mathcal{K}\mathbf{C}, \quad (12)$$

where  $\mathcal{K}$  is the  $M \times M$  dimensional positive definite matrix-valued kernel. Thus the cumbersome mapping problem boils down to finding the unique coefficient vector  $\mathbf{C}_{M \times D} = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_M\}^T$ . The energy function Eq. (9) can be rewritten as:

$$E(\mathbf{C}, \rho^2) = \frac{1}{2^d (\pi \rho)^{\frac{D}{2}}} + \mathcal{R}(\mathcal{T}) - \frac{2}{M} \sum_{m=1}^M \frac{1}{(2\pi \rho^2)^{\frac{D}{2}}}$$

$$\times \exp \left( -\frac{\|\hat{\mathbf{x}}_i - \mathbf{y}_i - \sum_{j=1}^M \mathcal{K}(\mathbf{y}_i, \mathbf{y}_j) \mathbf{c}_j\|^2}{2\rho^2} \right). \quad (13)$$

The kernel selection on  $\mathcal{K}$  of our method will be explained in the following part of this paper as it is strongly related to our global-local structural preservation.

### 3.3.2. Global-local structural preservation

Typically, non-rigid spatial transformations models can be categorized into two types according to their theoretical foundations [36]: (i) physically based models e.g., linear elasticity and fluid flow, and (ii) basis function expansions e.g., radial basis functions, B-splines and wavelets. Different types of models have their own constraints which reflect the prior knowledges. In our method, the Gaussian radial basis function is used to construct the kernel  $\mathcal{K}$ ,

with each entry  $\mathcal{K}(\mathbf{y}_i, \mathbf{y}_j) = \exp(-\frac{1}{2\rho^2} \|\mathbf{y}_i - \mathbf{y}_j\|^2)$ , and the constant  $\beta$  to control the spatial smoothness.

The regularization is formulated as:

$$\mathcal{R}(\mathcal{T}) = \lambda \cdot \mathcal{G}(\mathcal{T}) + \eta \cdot \mathcal{L}(\mathcal{T}), \quad (14)$$

where  $\mathcal{G}(\mathcal{T})$  is the global structure preservation (GSP) supported by the motion coherence theory (MCT) [23], and  $\mathcal{L}(\mathcal{T})$  is the local structure preservation (LSP) presented in this paper. The constant  $\lambda$  and  $\eta$  are the trade-off.  $\mathcal{G}$  and  $\mathcal{L}$  play complementary roles to improve the robustness and accuracy of transformation at both global and local scales.

**Global structure preservation (GSP):** In motion perception, there are a number of important phenomena involving coherence. Velocity coherence is a particular way of imposing smoothness on the underlying transformation. The concept of motion coherence is proposed in the motion coherence theory [23] which is intuitively interpreted as that points close to one another tend to move coherently. The GSP regularizes the norm of transformation to enforce the smoothness, and has the form:

$$\mathcal{G}(\mathcal{T}) = \|\mathcal{T}\|^2 = \mathcal{G}(\mathcal{V}) = \|\mathcal{V}\|^2 = 2\mathcal{K}\mathbf{C}. \quad (15)$$

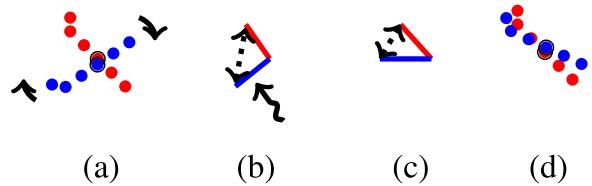
It is worth nothing that  $\mathcal{G}(\mathcal{T})$  and  $\mathcal{G}(\mathcal{V})$  are exactly the same, since the norm is invariant under affine transformation. It implies that motion coherence is imposed among the corresponding set  $\{\mathcal{T}(\mathbf{y}_i), \hat{\mathbf{x}}_i\}_{i=1}^M$ , and discouraging the undesired transformation, over the whole warping plane.<sup>2</sup> With a large  $\lambda$ , the constraint produces globally smooth transformation, while it produces more arbitrary transformation with small values.

**Local structure preservation (LSP):** Keeping the overall spatial relationship of the point set during the transformation is beneficial to preserving global structure. However, it may produce position deviations of the rest of the points when one point is mismatched, and may also be undesirable when source points need to be moved in different directions to match their target points at the same time [29]. In addition, at smaller stage of iterations, the local segment of  $\mathbf{x}_i$  and  $\mathbf{y}_j$ , i.e.,  $\{\hat{\mathbf{x}}_{ik}\}_{k=1}^K$  and  $\{\mathcal{T}(\mathbf{y}_{ik})\}_{k=1}^K$  are more likely to be dissimilar, consequently, the better alignment can be achieved which results from exerting "stronger force" over the segments other than the global constraint. To this end, we seek for minimizing the spatial deviations between the corresponding local segments  $\{\hat{\mathbf{x}}_{ik}\}_{k=1}^K$  and  $\{\mathcal{T}(\mathbf{y}_{ik})\}_{k=1}^K$ , and develop the LSP which takes the form:

$$\begin{aligned} \mathcal{L}(\mathcal{T}) &= \sum_{i=1}^M \left[ \sum_{k=1}^K w_{ik}^{\hat{\mathbf{x}}} (\hat{\mathbf{x}}_{ik} - \hat{\mathbf{x}}_i) - \sum_{k=1}^K w_{ik}^{\mathbf{y}} (\mathcal{T}(\mathbf{y}_{ik}) - \mathcal{T}(\mathbf{y}_i)) \right]^2 \\ &= \|(\mathbf{W}^{\hat{\mathbf{x}}} - \mathbf{K}\mathbf{I})\hat{\mathbf{X}} - (\mathbf{W}^{\mathbf{y}} - \mathbf{K}\mathbf{I})\mathcal{T}(\mathbf{Y})\|^2, \end{aligned} \quad (16)$$

where the two  $M \times M$  matrices  $\mathbf{W}^{\mathbf{y}}$  and  $\mathbf{W}^{\hat{\mathbf{x}}}$  denote the LSD weights for the source and putative target point sets, with each nonzero entry equaling to the respective weight for  $\hat{\mathbf{x}}_{ik}$  and  $\mathbf{y}_{ik}$ .  $K$  is the number of the KNN, and  $\mathbf{I}_{M \times M}$  is an identity matrix. The goal of the LSP can be interpreted to impose an alignment locally. An illustration of how the LSP works is shown in Fig. 3. Substituting the structure preservation terms into the energy function Eq. (13), we can therefore rewrite it in the complete form:

$$\begin{aligned} E(\mathbf{C}, \rho^2) &= \frac{1}{2^d(\pi\rho)^{\frac{d}{2}}} + \lambda \cdot \text{tr}(\mathbf{C}^T \mathcal{K} \mathbf{C}) + \eta \cdot \text{tr}(\mathbf{Q}^T \mathbf{Q}) \\ &\quad - \frac{2}{M} \sum_{i=1}^M \frac{1}{(2\pi\rho^2)^{\frac{d}{2}}} \end{aligned}$$



**Fig. 3.** The illustration of how the LSP works. Given a red segment and its rotated form drawn in blue in Fig. 3(a), the colored lines in Fig. 3(b) and Fig. 3(c) denote the LSD of the segments, respectively. The snake arrow in Fig. 3(b) denotes the imaginary force exerted on the LSD of the blue segment, and the dotted black lines denote the local discrepancy between two segments. Two segments are aligned in Fig. 3(d). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\times \exp \left( -\frac{\|\hat{\mathbf{x}}_i - \mathbf{y}_i - \sum_{j=1}^M \mathcal{K}(\mathbf{y}_i, \mathbf{y}_j) \mathbf{c}_j\|^2}{2\rho^2} \right), \quad (17)$$

where  $\mathbf{Q} = (\mathbf{W}^{\hat{\mathbf{x}}} - \mathbf{K}\mathbf{I})\hat{\mathbf{X}} - (\mathbf{W}^{\mathbf{y}} - \mathbf{K}\mathbf{I})\mathcal{T}(\mathbf{Y})$ . Finally, the transformation is obtained by taking the derivative of the above energy function Eq. (17) with respect to the coefficient matrix  $\mathbf{C}$ , which is given by:

$$\frac{\partial(E)}{\partial(\mathbf{C})} = \frac{2\mathcal{K}[\mathbf{U} \odot (\mathbf{V} \otimes \mathbf{1}_{1 \times d}) + \mathbf{B}^T \mathbf{BC} - \mathbf{B}^T \mathbf{A}]}{M\rho^2(2\pi\rho^2)^{\frac{d}{2}}} + 2\lambda\mathcal{K}\mathbf{C}, \quad (18)$$

where  $\mathbf{U}_{M \times D} = \mathbf{Y} + \mathcal{K}\mathbf{C} - \hat{\mathbf{X}}$ ,  $\mathbf{V}_{M \times 1} = \exp(\text{diag}(\mathbf{U}\mathbf{U}^T)/2\rho^2)$ ,  $\mathbf{B}_{M \times M} = \mathbf{W}^{\mathbf{y}} - \mathbf{K}\mathbf{I}$  and  $\mathbf{A}_{M \times D} = (\mathbf{W}^{\hat{\mathbf{x}}} - \mathbf{K}\mathbf{I})\hat{\mathbf{X}} - (\mathbf{W}^{\mathbf{y}} - \mathbf{K}\mathbf{I})\mathbf{Y}$ . Symbols  $\odot$  and  $\otimes$  denote the Hadamard product and the Kronecker product, respectively. The symbol  $\text{diag}(\cdot)$  is the diagonal of a matrix, and  $\text{tr}(\cdot)$  is the trace.

A gradient-based numerical optimization technique is employed to solve this optimization problem. Since the non-rigid point set registration has high degrees of flexibility, trapping the optimization process by local minima, the convergence is improved by adopting deterministic annealing with a manner of coarse-to-fine on the covariances  $\rho^2$ . It initializes with a large value for  $\rho^2$  which progressively decreases by  $\rho^2 \leftarrow \varepsilon\rho^2$ , where  $\varepsilon$  is a constant. And the energy function Eq. (17) will converge to an optimal solution.

### 3.4. Fast implementation and regularization softening

We have explained the main process, the major contributions on the CE and TE steps of our method. In addition to the aforementioned details, the computational complexity and registration accuracy are further optimized by using (i) sparse approximation and (ii) regularization softening.

**Sparse approximation:** The kernel matrix  $\mathcal{K}$  plays a crucial role in the formulation of the  $L_2E$  criterion and the regularization. However, it contains  $M \times M$  entries which causes significant complexity problems when handling large amount of points. Several well-studied techniques such as the low-rank matrix approximation [37] and the subset of regressors method [38,39] can be used to address this problem. In this paper, the second one is used by randomly picking only a subset of size  $M'$  ( $M' \ll M$ ) source points  $\tilde{\mathbf{Y}} = \{\tilde{\mathbf{y}}_j | \tilde{\mathbf{y}}_j \in \mathbb{R}^D, j = 1, \dots, M'\}$ , which means that the global and local regularizations are both limited within the correspondences subset  $\{\tilde{\mathbf{y}}_i, \hat{\mathbf{x}}_i\}_{i=1}^{M'}$ . Hence, the original  $M \times M$  dimensional kernel  $\mathcal{K}$  can be reduced to a  $M' \times M'$  dimensional matrix  $\tilde{\mathcal{K}}$ , where each entry  $\tilde{\mathcal{K}}(\tilde{\mathbf{y}}_i, \tilde{\mathbf{y}}_j) = \exp(-\frac{1}{2\rho^2} \|\tilde{\mathbf{y}}_i - \tilde{\mathbf{y}}_j\|^2)$ . The solution presented in Theorem 1 now has the form:

$$\mathcal{V}(\mathbf{y}_i) = \sum_{j=1}^{M'} \mathcal{K}(\mathbf{y}_i, \tilde{\mathbf{y}}_j) \tilde{\mathbf{c}}_j. \quad (19)$$

<sup>2</sup> Do not confuse  $\{\mathbf{y}_i, \hat{\mathbf{x}}_i\}_{i=1}^M$  and  $\{\mathcal{T}(\mathbf{y}_i), \hat{\mathbf{x}}_i\}_{i=1}^M$ , where the former is the correspondence estimated in the CE step, and the latter is its transformed counterpart in the TE step.

Substituting the sparse kernel and Eq. (19) into the energy function Eq. (17), it can be rewritten it as:

$$E(\tilde{\mathbf{C}}, \rho^2) = \frac{1}{2^d (\pi \rho)^{\frac{d}{2}}} + \lambda \cdot \text{tr}(\tilde{\mathbf{C}}^T \tilde{\mathbf{K}} \tilde{\mathbf{C}}) + \eta \cdot \text{tr}(\tilde{\mathbf{Q}}^T \tilde{\mathbf{Q}}) - \frac{2}{M} \sum_{i=1}^M \frac{1}{(2\pi \rho^2)^{\frac{d}{2}}} \exp\left(-\frac{\|\tilde{\mathbf{x}}_i - \mathbf{y}_i - \Gamma_{(i,\cdot)} \tilde{\mathbf{C}}\|^2}{2\rho^2}\right), \quad (20)$$

where  $\tilde{\mathbf{C}}_{M' \times D} = \{\tilde{\mathbf{c}}_1, \tilde{\mathbf{c}}_2, \dots, \tilde{\mathbf{c}}_{M'}\}^T$ ,  $\Gamma$  is a  $M \times M'$  basis matrix with each entry  $\Gamma_{ij} = \mathcal{K}(\mathbf{y}_i, \tilde{\mathbf{y}}_j)$ ,  $\Gamma_{(i,\cdot)}$  denotes the  $i$ th row of  $\Gamma$ . And  $\tilde{\mathbf{Q}} = (\tilde{\mathbf{W}}^{\tilde{\mathbf{x}}} - \mathbf{K}\mathbf{I})\tilde{\mathbf{X}} - (\tilde{\mathbf{W}}^{\mathbf{y}} - \mathbf{K}\mathbf{I})\mathcal{T}(\tilde{\mathbf{Y}})$ , where  $\mathbf{I}_{M' \times M'}$  is an identity matrix. Note that all the components of  $\tilde{\mathbf{Q}}$  follows the same rule for obtaining  $\mathbf{Q}$  specifically, each point and its KNNs are both found from the correspondences subset  $\{\tilde{\mathbf{y}}_i, \tilde{\mathbf{x}}_i\}_{i=1}^{M'}$ . Taking the derivative of Eq. (20) with respective to  $\tilde{\mathbf{C}}$  we obtain:

$$\frac{\partial(\tilde{E})}{\partial(\tilde{\mathbf{C}})} = \frac{2\Gamma^T [\tilde{\mathbf{U}} \odot (\tilde{\mathbf{V}} \otimes \mathbf{1}_{1 \times d}) + \tilde{\mathbf{B}}^T \tilde{\mathbf{B}} \tilde{\mathbf{C}} - \tilde{\mathbf{B}}^T \tilde{\mathbf{A}}]}{M\rho^2(2\pi\rho^2)^{\frac{d}{2}}} + 2\lambda \tilde{\mathbf{K}} \tilde{\mathbf{C}}, \quad (21)$$

where  $\tilde{\mathbf{U}}_{M \times D} = \mathbf{Y} + \Gamma \tilde{\mathbf{C}} - \tilde{\mathbf{X}}$ ,  $\tilde{\mathbf{V}}_{M \times 1} = \exp(\text{diag}(\tilde{\mathbf{U}} \tilde{\mathbf{U}}^T)/2\rho^2)$ ,  $\tilde{\mathbf{B}}_{M' \times M'} = \tilde{\mathbf{W}}^{\mathbf{y}} - \mathbf{K}\mathbf{I}$  and  $\tilde{\mathbf{A}}_{M' \times D} = (\tilde{\mathbf{W}}^{\tilde{\mathbf{x}}} - \mathbf{K}\mathbf{I})\tilde{\mathbf{X}} - (\tilde{\mathbf{W}}^{\mathbf{y}} - \mathbf{K}\mathbf{I})\tilde{\mathbf{Y}}$ .

**Regularization softening:** Unlike most of the iterative non-rigid point set registration methods we introduced in Section 1, ICP [14] writes the energy function in the form without regularization. It performs well in rigid case since the mappings are strictly preserved. Inspired by ICP, the energy functions Eqs. (17) and (20) are transformed to an ICP-like form by progressively reducing the strength of the preservation terms. The function  $\tau = (t_\sigma^4 - t^4 + 1)^{1/4}/t_\sigma$  is formulated to determine the softening parameter  $\tau$ , where  $t_\sigma$  is the maximum iteration number.  $\tau$  remains stable in most iterations, then sharply decreasing to zero in the last several iterations. The experiment shows that the softened  $\lambda = \tau\lambda$  and  $\eta = \tau\eta$  improve the performances of our method. The pseudo-code of point sets registration using our method is outlined in Algorithm 1.

### 3.5. Computational complexity

Recall that we have  $M$  source points  $\{\mathbf{y}_j\}_{j=1}^M$ ,  $N$  target points  $\{\mathbf{x}_i\}_{i=1}^N$  and  $M$  putative target points  $\{\tilde{\mathbf{x}}_i\}_{i=1}^M$ . In the CE step, a k-d tree structure can determine the KNNs efficiently, since  $K \ll M$  and  $K \ll N$ , the KNNs are obtained by  $K$  operations of sequential search with  $\mathcal{O}(KM)$  or  $\mathcal{O}(KN)$  complexity for one point. The posterior probability matrix  $\mathbf{P}$  requires  $\mathcal{O}(NM)$  work to compute. In the TE step, the weight matrix  $\mathbf{W}$  is of  $M$  rows, totally  $K$  nonzero entries each of which require  $K$  iterations to compute. Hence, it requires  $\mathcal{O}(K^2M)$  work to compute. The Matlab Optimization toolbox, e.g., the Matlab function fminunc, which implicitly uses the Broyden-Fletcher-Goldfarb-Shanno (BFGS) Quasi-Newton method with a mixed quadratic and cubic line search procedure, is used for solving the numerical optimization method, and the total complexity is approximately  $\mathcal{O}(M^3 + 2M^2)$ . Overall, the computational complexity of our method is  $\mathcal{O}(M^3)$ . In the sparse approximation,  $M'$ , the number of the randomly selected points required to construct the transformation in Eq. (19), is in general not large.  $M'$  is set to 35 for all the results. Therefore, the complexity is reduced to  $\mathcal{O}(M'^3 + M'^2 + M'M)$  which is close to  $\mathcal{O}(M)$  since  $M' \ll M$ .

## 4. Experiment

Our method is tested using three types of data: (i) synthetic data, (ii) sequence data, and (iii) real data against three kinds of state-of-the-art methods (totally nine methods are included):

---

**Algorithm 1:** Non-rigid point set registration using DFMM and GLSP.

---

```

input : Two point sets  $\mathbf{X}$  and  $\mathbf{Y}$ 
output : Transformed point set  $\hat{\mathbf{Y}}$ 
parameter:  $r, \omega, \beta, \epsilon, \varepsilon, \lambda, \eta$  and  $\tau$ 
1 Initialize the DFMM covariances  $\sigma^2$ ;
2 Construct Gaussian kernel  $\mathcal{K}$ ;
3 repeat
4   Correspondence estimation:
5   | Compute the LSDs and MSDs of  $\mathbf{X}$  and  $\mathbf{Y}$ ;
6   | Compute the posterior probability matrix  $\mathbf{P}$  by Eq. (7);
7   | Compute the putative target point set by  $\hat{\mathbf{X}} = \mathbf{P}\mathbf{X}$ ;
8   end
9   Transformation estimation:
10  | Initialize the Gaussian covariances  $\rho^2$  and the LSD weight  $\mathbf{W}$ ;
11  | repeat
12  |   | Optimize the energy function Eq. (17) by a numerical
13  |   | technique using the gradient function Eq. (18);
14  |   | Update the coefficient matrix  $\mathbf{C}$ ;
15  |   | Anneal  $\rho^2 \leftarrow \varepsilon\rho^2$ ;
16  |   until reach  $t_\rho$  iteration number;
17  |   Update the source point set by  $\mathbf{Y} \leftarrow \mathcal{T}(\mathbf{Y})$ ;
18  end
19  | Anneal  $\sigma^2 \leftarrow \epsilon\sigma^2$ ;
20 until reach  $t_\sigma$  iteration number;
21 The transformed source point set  $\hat{\mathbf{Y}}$  is obtained in the final iteration.

```

---

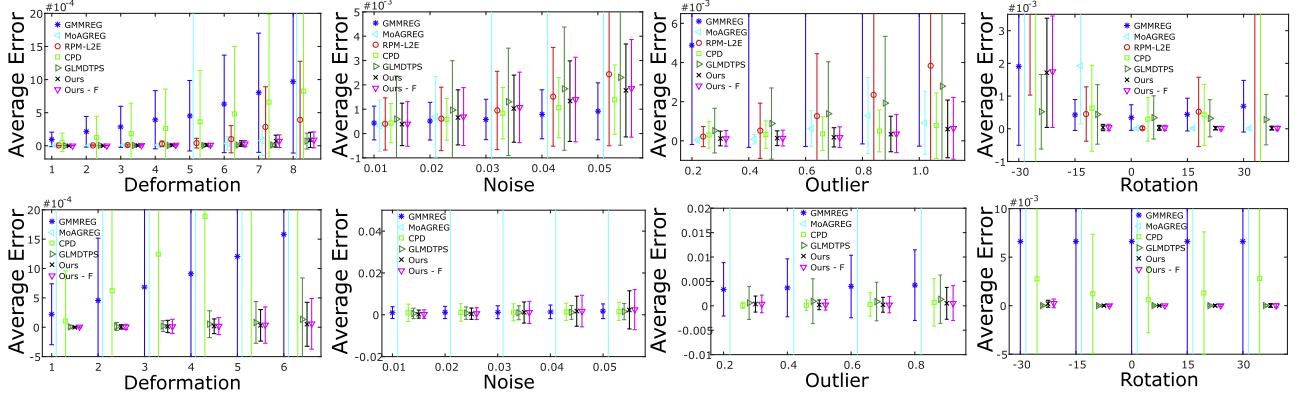
- (i) Iterative methods: GMMREG [20], MoAGREG [30], RPM-L2E [26], CPD [22], GLMDTPS [29] and Nicholas et al. [21].
- (ii) Non-learning graph based methods: FGM [40].
- (iii) Learning graph based methods: Caetano et al. [41] and Leordeant et al. [42].

The performances of our method by using softened  $\lambda$  and  $\eta$  against the constant counterpart are also tested. Two criteria are used to evaluate the performances of every method. (i) Root mean square error (RMSE):  $RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^M |\mathbf{x}_i - \hat{\mathbf{y}}_i|^2}$ , where  $|\mathbf{x}_i - \hat{\mathbf{y}}_i|$  is the Euclidean distance. (ii) Recall:  $Recall = TP/(TP + FN)$ , where  $TP$  denotes true-positive and  $FN$  denotes false-negative. All experiments are implemented in Matlab R2016a on a Intel Core i7-6700HQ CPU 2.60 GHz with 16GB RAM.

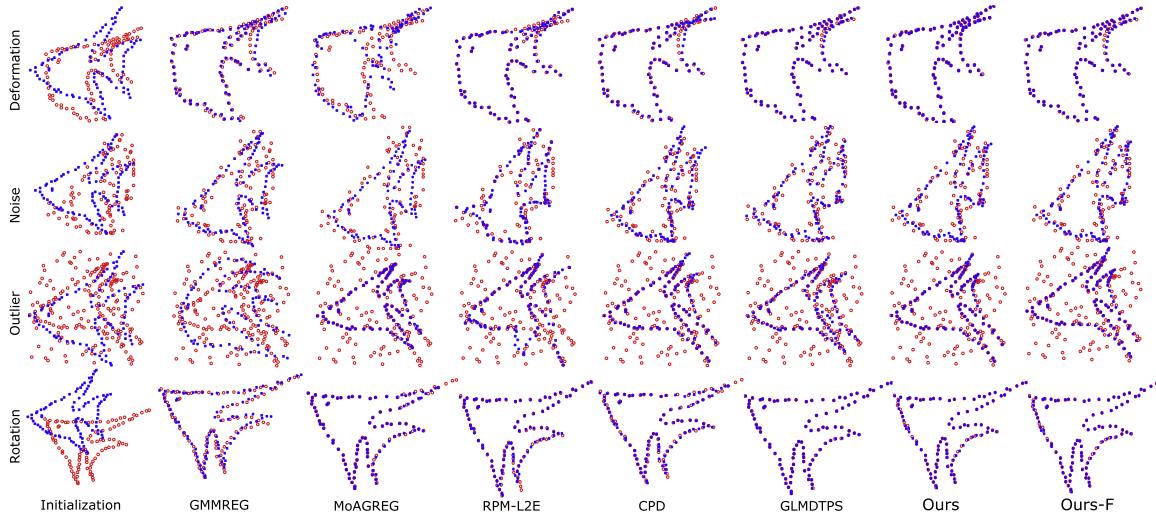
### 4.1. Parametric setting

Seven groups of free parameters are included: the iteration number  $t_\sigma$  and  $t_\rho$ , the annealing parameters  $r, \epsilon, \varepsilon$  and  $\tau$ , the outlier weight  $\omega$ , covariances  $\sigma^2$  and  $\rho^2$ , the regularization parameters  $\lambda$  and  $\eta$ , the smooth controller  $\beta$ , and the number of neighbors  $K$ . Default settings are followed according to the respective literatures for all compared methods.

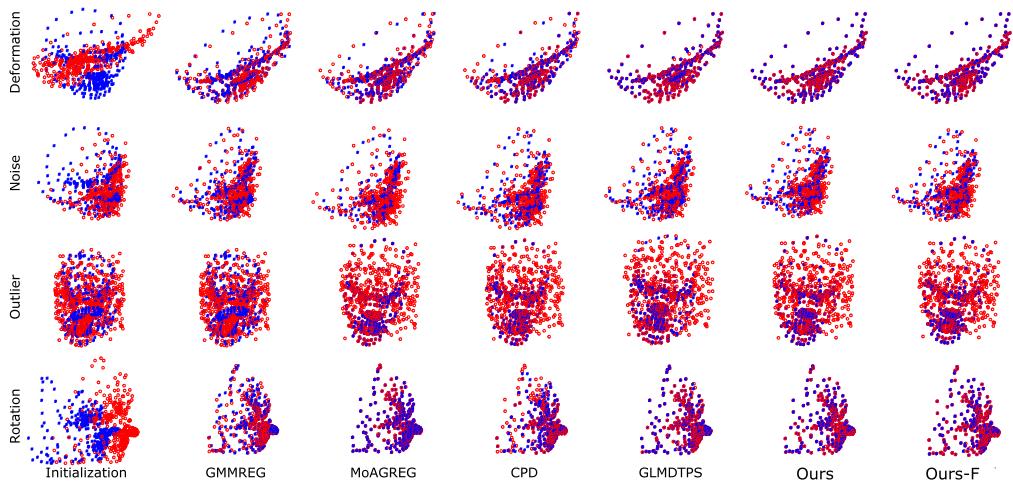
- Iteration control: According to an initial trial-and-error experiment using the fish [15,26,43], the iteration number for our method to converge is set to  $t_\sigma = 100$ . The Quasi-Newton method is implemented by using the function fminunc in Matlab with the termination condition:  $t_\rho = 50$ .
- Annealing parameters: There are four annealing parameters.  $r$  determines the trade-off for constructing the MSDs using the point-wise Euclidean distances and LSDs.  $r$  is set to 8 so that at the beginning of iteration,  $\alpha = \exp(-0/8) = 1$ , and at the



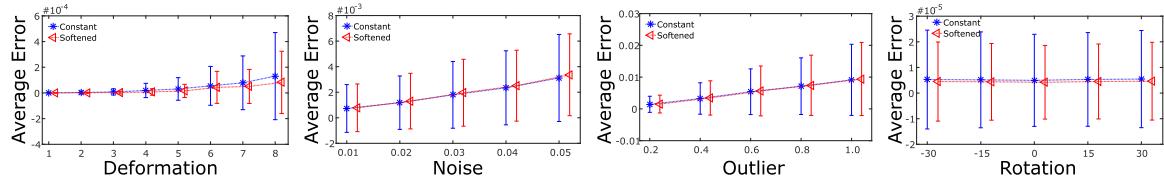
**Fig. 4.** Registration performances of our method on the *fish* (upper row) and 3D face point set (lower row). For the *fish*, our method ( $\times$ ) and its fast version ( $\nabla$ ) compared with GMMREG (\*), MoAGREG ( $\triangleleft$ ), RPM-L2E ( $\circ$ ), CPD ( $\square$ ) and GLMDTPS ( $\triangleright$ ). For the 3D face, RPM-L2E is excluded since its published version focuses on the 2D case. The error bars indicate the standard deviations of the mean errors in 100 random experiments.



**Fig. 5.** Registration examples on the *fish*. Rows from top to bottom are examples in the deformation (the 8th degree), noise (0.05), outlier (1.0) and rotation (30°) experiments. The source point sets and target point sets here are presented by blue \* and red  $\circ$ , respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 6.** Registration examples on the 3D face. Rows from top to bottom are examples in deformation (the 6th degree), noise (0.05), outlier (0.8) and rotation (30°) experiments. The source point sets and target point sets here are presented by blue \* and red  $\circ$ , respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 7.** Comparison of softened regularization (\*) against its constant counterpart (<). The error bars indicate the standard deviations of the mean errors in 100 random experiments.

final iteration,  $\alpha = \exp(-100/8) \approx 3 \times 10^{-6}$ , by which the interchange of the global and local feature is realized.  $\epsilon$  and  $\varepsilon$  are for controlling the rate of deterministic annealing over the covariances  $\sigma^2$  and  $\rho^2$ , respectively. It can be set closely to 1 for acquiring more satisfied result, yet considering the efficiency issue,  $\epsilon$  and  $\varepsilon$  are set to 0.9 and 0.5 respectively. The last one  $\tau$  softens the strength of the two regularization parameters  $\lambda$  and  $\eta$ . Hence,  $\tau$  is set by a well defined function  $\tau = (t_\sigma^4 - t^4 + 1)^{1/4}/t_\sigma^4$ , where  $t$  and  $t_\sigma$  denote the current and maximum iteration number, respectively.

- Outlier weighting:  $\omega$  is set to 0.2 for deformation and rotation experiments, and to 0.7 for noise and outlier experiments.
- Covariances:  $\sigma^2$  determines the searching range of the DFMM and is initialized to 1,  $\rho^2$  controls the convergence of the  $L_2E$  criterion based energy function and is initialized to 0.05.
- Regularization parameters: Two regularization parameters  $\lambda$  and  $\eta$  are included in our method. They represent the weights of the global-local structural preservation term, respectively.  $\lambda$  and  $\eta$  are both set to 2.
- Smooth controller:  $\beta$  flexibly determines the locality of spatial smoothness, thus,  $\beta$  is set to 2.
- The number of neighbor points: In order to distinguish between a corner and a cross, at least four neighboring points are required [29]. Thus  $K$  is set to 5.

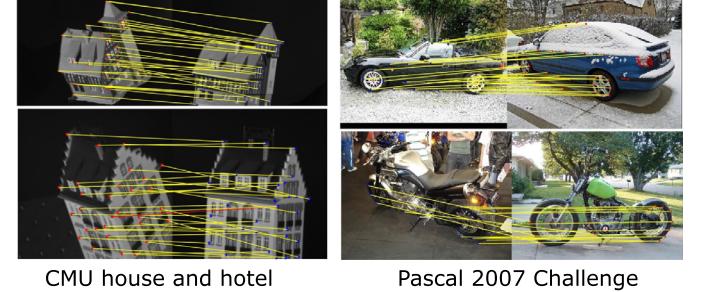
#### 4.2. Result on synthetic data

In the first series of experiments, the performances of our method and its fast version (Ours-F) are compared with five state-of-the-art methods: GMMREG [20], MoAGREG [30], RPM-L2E [26], CPD [22] and GLMDTPS [29] using the synthesized datasets [15,26,29,43,44]. The fish contains 98 points and the 3D face contains 392 points. The experiments are designed following the previous work [29].

The performances (the RMSE and STD) of all the methods on the fish and 3D face point sets are shown in Fig. 4. Since the published version of RPM-L2E focuses on the 2D case, it is excluded from the comparison on 3D face. Our method achieves the best performances in outlier, 3D deformation and over degrees from 2 to 5 of rotation scenarios. Of all the scenarios, our method performs slightly better than its fast version. In the noise experiment, all the methods except MoAGREG give accurate alignments while GMMREG generally performs better. Registration examples in 2D and 3D cases are shown in Fig. 5 and Fig. 6, respectively.

#### 4.3. Constant vs. softened

To demonstrate the availability of the softened regularization, the performances of our method using softened regularization are compared against its constant counterpart. All the settings are the same as in Section 4.2. The results shows that obvious improvements are achieved over all degrees of deformation and rotation, slightly worse performances are presented in the outlier scenario, as shown in Fig. 7.



**Fig. 8.** Matching examples on CMU sequence dataset (left) and Pascal 2007 challenge (right). The correct matches are shown by the yellow lines and the mismatches are shown by the red lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

#### 4.4. Result on sequence data

The CMU sequence data is one of the most popular datasets and has been widely used to test the performances of registration methods. The CMU house and hotel dataset consist of 111 and 101 frames, respectively, all of which have 30 labeled landmarks per frame. The performances of our method with six state-of-the-art methods: GMMREG [20], RPM-L2E [26], CPD [22], GLMDTPS [29], Caetano et al. [41] and Leordeanu et al. [42] are compared under all possible data pairs. The results on the matching rate are shown in Table 3. Our method gives the near-perfect matching results in all possible data pairs. Two representative matching examples are shown in Fig. 8.

#### 4.5. Result on real data

In the third series of experiments, the performances of our method are tested in three types of datasets: (i) a dataset from [29], it consists of 30 pairs of car images and 20 pairs of motor-bike images selected from Pascal 2007 Challenge, each pair contains 30–60 feature points; (ii) a pair of CT images and two pairs of MRI images from [45], the point sets are created by the edges of the objects and extracted using the Matlab function *imcontour*; (iii) five pairs of retinal images in multi-modality scenarios and another five in multi-view scenarios; (iv) two pairs of remote sensing images (Washington and Las Vegas) captured from Google Earth. For dataset (iii) and (iv), the SIFT algorithm [46] is used to extract feature points, and all feature points in dataset (iii) are preprocessed following the method proposed in [47].

##### 4.5.1. Pascal 2007 challenge dataset

The performances of our method are compared against GMMREG [20], MoAGREG [30], RPM-L2E [26], CPD [22], FGM [40], and Leordeanu et al. [42]. The results on the matching rate are shown in Table 4. Our matching rate is higher than the published results of other methods. Some matching examples are shown in Fig. 8.

**Table 3**

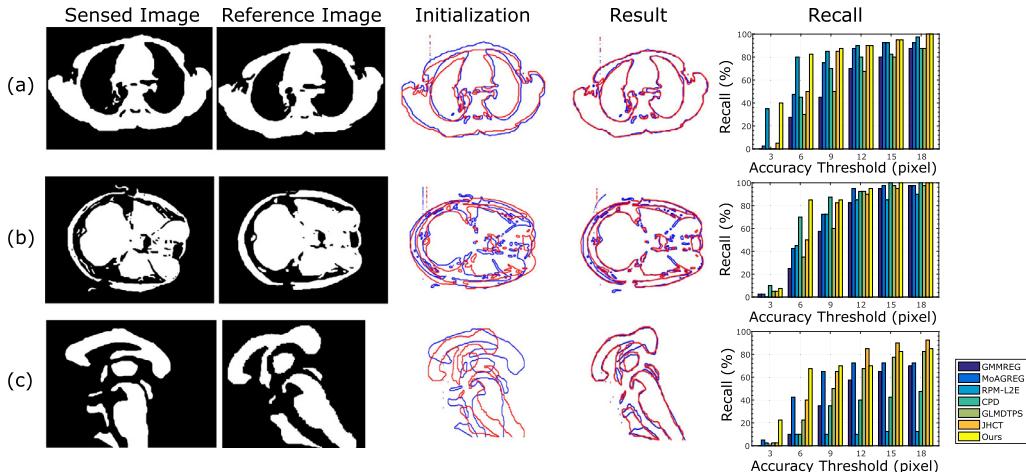
The matching rates on CMU sequence datasets. CPD, GMMREG, RPM-L2E and our method are tested. For Caetano et al. and Leordeanu et al., their published results are reported. S and U denote ‘supervised’ and ‘unsupervised’, respectively. The numbers in S and U denote the number of training image pairs.

	GMMREG	RPM-L2E	CPD	GLMDTPS	Caetano et al.	Leordeanu et al.	Ours
Setting					S(106)	S(5)	U(5)
House	99.5%	99.4%	99.6%	100%	< 96%	99.8%	99.8%
Hotel	97.1%	98.06%	98.9%	90.77%	< 90%	94.8%	94.8%

**Table 4**

The matching rate on Pascal 2007 Challenge dataset. GMMREG, MoAGREG, RPM-L2E, CPD and our method are tested. For FGM and Leordeanu et al., their published results are reported. L: after learning.

	GMMREG	MoAGREG	RPM-L2E	CPD	FGM	Leordeanu et al.	Ours
Pascal 2007	82.9%	76.1%	82.6%	83.5%	80%	80%	89.4%



**Fig. 9.** Registration examples and performances on medical dataset. The recall are used to evaluate on the three real data experiments under the accuracy threshold 3–18 pixels. Rows from top to bottom are examples and performances on: thorax CT data, transverse plane brain MRI data, and sagittal plane brain MRI data, respectively. The units of the accuracy thresholds are in pixels.

**Table 5**

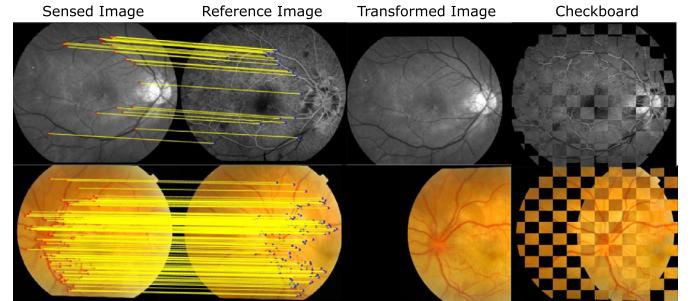
The RSME on retinal images. The units of all data are in pixels.

GMMREG	MoAGREG	RPM-L2E	CPD	GLMDTPS	Ours
$4.73 \pm 3.84$	$5.45 \pm 2.49$	$4.49 \pm 0.15$	$4.45 \pm 2.61$	$4.38 \pm 0.92$	$3.76 \pm 0.93$

#### 4.5.2. Medical dataset

**CT and MRI data:** the performances of our method are compared against GMMREG [20], MoAGREG [30], RPM-L2E [26], and CPD [22], and Nicholas et al. [21] (abbreviated as JHCT for convenience). For methods except JHCT, 40 landmarks are selected from the contours of the sensed images and reference images, respectively. The recalls are obtained by evaluating the registration results based on the landmark sets under different pixel thresholds. For JHCT, since it is intensity-based, its landmarks are selected from the sensed images and transformed images. Each image has the resolution of  $615 \times 460$ . The quantitative results and the matching examples of our method on the whole contour point sets are shown in Fig. 9. We see that JHCT and our method generally outperform other methods where JHCT is relatively better. This is reasonable because JHCT is designed for image registration, therefore the intensity information can be exploited.

**Retinal data:** the performances of our method are compared against GMMREG [20], MoAGREG [30], RPM-L2E [26], CPD [22] and GLMDTPS [29]. Since retinal images are obtained by a fundus camera and may be taken from different perspectives or with different modalities, accuracy registration is important for acquiring comprehensive understanding and diagnosing retinal disease. The

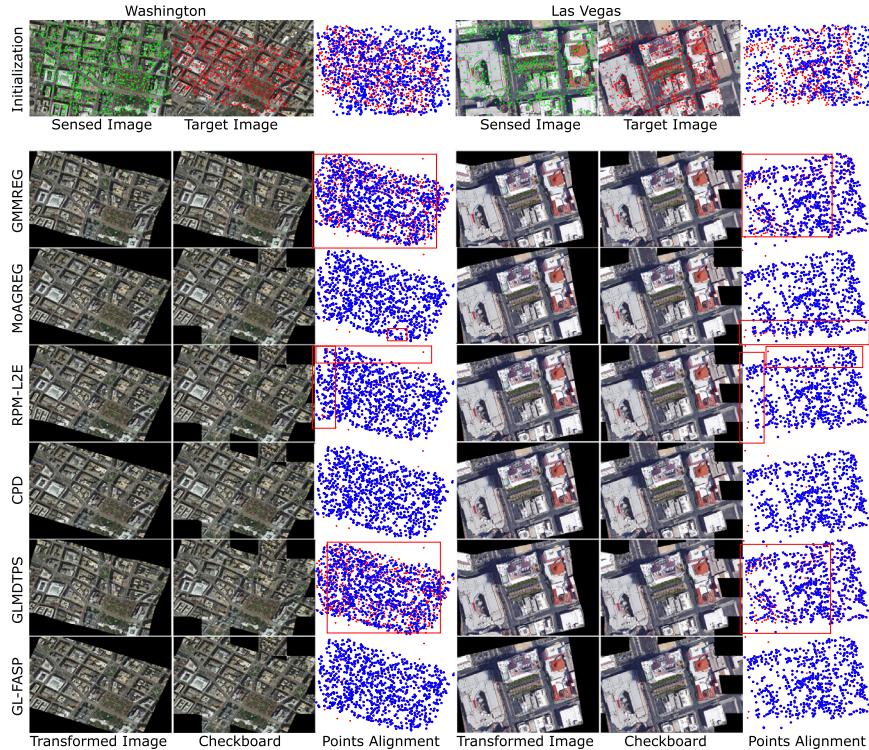


**Fig. 10.** Registration results of our Ours on retinal images. Rows from top to bottom are results on multi-modality scenario and multi-view scenario, respectively. All feature points are extracted follow the method proposed in [47] and correctly registered.

quantitative comparison using RMSE, as well as two representative matching results are shown in Table 4 and Fig. 10, respectively.

#### 4.5.3. Remote sensing images

In this section, experiments on satellite images obtained from Google Earth are conducted. The dataset consists of two image pairs captured from Washington and Las Vegas. Each image has the resolution of  $586 \times 452$  and contains imaging viewpoint changes. The performances of our method are compared against GMMREG [20], MoAGREG [30], RPM-L2E [26], CPD [22] and GLMDTPS [29]. The registration results and the demonstration on feature points alignment are shown in Fig. 11. At least 10 pairs of corresponding



**Fig. 11.** Registration results of remote sensing images. Top row: two pairs of remote sensing images which are captured from Washington (left) and Las Vegas (right). Green and red dots: the SIFT feature points extracted from the sensed and target images, respectively. Rows from second to last are the registration results using GMMREG, MoAGREG, RPM-L2E, CPD, GLMDTPS and our method. The transformed image, the  $5 \times 5$  checkboard and feature points alignment are shown within each row. Some of the errors are highlighted by red rectangles. We see that large errors presented on the alignment of the feature points lead to undesired distortion in the transformed images. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

**Table 6**  
The RMSE on Remote sensing dataset. The units of all data are in pixels.

	GMMREG	MoAGREG	RPM-L2E	CPD	GLMDTPS	Ours
Washington	13.52	0.99	4.08	0.9	6.33	0.85
Las Vegas	12.23	1.18	1.66	1.57	5.97	1.08

points between the transformed images and the reference images are manually determined for quantitative comparisons, note that all corresponding points are well-distributed and selected on the easily identified places. The registration error is defined using the RMSE, as shown in Table 6.

## 5. Conclusion

We have presented a dual-feature based point set registration method with global-local correspondence and transformation estimations. The main idea of our method is to first estimate the correspondence via the DFMM, and then update the transformation under the constraint of global-local structural preservation terms. Compared with the current methods, the major differences and advantages of this work include: (i) the dual-feature finite mixture model (DFMM) is constructed, with its searching ranges directly controlled by a deterministic scheme; (ii) the vector-based local structure descriptor (LSD) is defined, together with the Euclidean coordinate it can project point sets to the mixture structure descriptor (MSD) space; (iii) The LSD is then incorporated into our DFMM, by which the robustness and accuracy on correspondence estimation is improved; (iv) the local structure preservation (LSP) which preserves the local structure of the point set during the transformation is formulated, and two structurally complementary (global/local) constraints are employed for regularizing the  $L_2 E$

based energy function; (v) our method is robust to outlier contamination since the DFMM and the  $L_2 E$  estimator are included; (vi) the energy function of our method is also sharply simplified in the last several iterations, leading to the maximization of the point-wise alignment by releasing the constraints. Experiments on synthetic data, sequence data and real data are undertaken. Our method shows favorable performances when compared with nine state-of-the-art methods.

## Conflict of interest

The authors declare that they have no conflicts of interest in the research.

## Acknowledgments

The authors thank Jiayi Ma, Andriy Myronenko, Xubo Song, Bing Jian, Baba C. Vemuri, Feng Zhou, Gang Wang, Yang Yang, Marius Leordeanu and Tiberio S. Caetano for providing their implementation source codes and test datasets. This greatly facilitated the comparison experiments. This work was supported by (i) National Nature Science Foundation of China (41661080); (ii) Key Scientific Research Project of Education Department of Yunnan Province (2015Z069); (iii) Scientific Research Foundation of Yunnan Provincial Department of Education (2017TYS045); (iv) Doctoral Scientific Research Foundation of Yunnan Normal University (01000205020503065); (v) Student Research Training Program of Yunnan Normal University (ky2016-114).

## References

- [1] G. Tam, Z. Cheng, Y. Lai, F. Langbein, A.D. Marshall, A.D. Marshall, R. Martin, X. Sun, P. Rosin, Registration of 3d point clouds and meshes: a survey from rigid to nonrigid, IEEE Trans. Vis. Comput. Graph. 19 (7) (2013) 1199–1217.

- [2] B. Zitova, J. Flusser, Image registration methods: a survey, *Imag. Vis. Comput.* 21 (11) (2003) 977–1000.
- [3] J. Ma, J. Zhao, Y. Ma, J. Tian, Non-rigid visible and infrared face registration via regularized gaussian fields criterion, *Pattern Recognit.* 48 (3) (2015) 772–784.
- [4] Z. Wei, Y. Han, M. Li, K. Yang, Y. Yang, Y. Luo, S. Ong, A small UAV based multi-temporal image registration for dynamic agricultural terrace monitoring, *Remote Sens.* 9 (9) (2017) 904.
- [5] S.K. Yong, J.H. Lee, J.B. Ra, Multi-sensor image registration based on intensity and edge orientation information, *Pattern Recognit.* 41 (11) (2008) 3356–3365.
- [6] J. Han, B. Bhanu, Fusion of color and infrared video for moving human detection, *Pattern Recognit.* 40 (6) (2007) 1771–1784.
- [7] F. Colas, R. Siegwart, A Review of Point Cloud Registration Algorithms for Mobile Robotics, Now Publishers Inc., 2015.
- [8] A. Sotiras, C. Davatzikos, N. Paragios, Deformable medical image registration: a survey, *IEEE Trans. Med. Imag.* 32 (7) (2013) 1153–1190.
- [9] W.S. Ronald, T.W. Tang, C. Albert, Non-rigid image registration of brain magnetic resonance images using graph-cuts, *Pattern Recognit.* 44 (10–11) (2011) 2450–2467.
- [10] K. Yang, G. Huangfu, W. Rao, M. Sun, S. Zhang, Y. Luo, Y. Yang, Quake warning funds on shaky ground, *Science* 358 (6368) (2017). 1263–1263.
- [11] K. Yang, Z. Yu, Y. Luo, Y. Yang, L. Zhao, X. Zhou, Spatial and temporal variations in the relationship between lake water surface temperatures and water quality - A case study of Dianchi lake, *Sci. Total Environ.* 624 (2018) 859–871.
- [12] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, J. Tian, Robust feature matching for remote sensing image registration via locally linear transforming, *IEEE Trans. Geosci. Remote Sens.* 53 (12) (2015) 6469–6481.
- [13] K. Yang, A. Pan, Y. Yang, S. Zhang, S. Ong, H. Tang, Remote sensing image registration using multiple image features, *Remote Sens.* 9 (6) (2017) 581.
- [14] P. Besl, N. McKay, A method for registration of 3d shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (1992) 239–256.
- [15] H. Chui, A. Rangarajan, A new algorithm for non-rigid point matching, *Comput. Vis. Image Underst.* 89 (2003) 114–141.
- [16] H. Chui, J. Rambo, R. Duncan, A. Rangarajan, Registration of cortical anatomical structures via robust 3d point matching, in: *Information Processing in Medical Imaging*, 1999, pp. 168–181.
- [17] A. Yuille, J. Kosowsky, Statistical physics algorithms that converge, *Neural Comput.* 6 (1994) 341–356.
- [18] F. Bookstein, Principal warps: thin-plate splines and the decomposition of deformations, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989) 567–585.
- [19] A. Tikhonov, V. Arsenin, *Solutions of Ill-Posed Problems*, Winston and Sons, Washington, D.C., 1977.
- [20] B. Jian, B. Vemuri, Robust point set registration using gaussian mixture models, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (2011) 1633–1645.
- [21] J. Nicholas, P. Suyash, S. Gang, S. Tessa, C. James, Point set registration using Havrda-Charvat-Tsallis entropy measures, *IEEE Trans. Med. Imag.* 30 (2011) 451–460.
- [22] A. Myronenko, X. Song, Point set registration: coherent point drift, *IEEE Trans. Pattern Anal. Mach. Intell.* 32 (2010) 2262–2275.
- [23] A. Yuille, N. Grzywacz, A mathematical analysis of the motion coherence theory, *Int. J. Comput. Vis.* 3 (1989) 155–175.
- [24] L. Greengard, J. Strain, The fast Gauss transform, *J. Sci. Stat. Comput.* 12 (1991) 79–94.
- [25] I. Markovsky, Structured low-rank approximation and its applications, *Automatica* 44 (2008) 891–909.
- [26] J. Ma, W. Qiu, J. Zhao, Y. Ma, Robust  $L_2$ E estimation of transformation for non-rigid registration, *IEEE Trans. Signal Proc.* 63 (2015) 1115–1129.
- [27] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (2002) 509–521.
- [28] A. Basu, I. Harris, N. Hjort, M. Jones, Robust and efficient estimation by minimising a density power divergence, *Biometrika* 85 (1998) 549–559.
- [29] Y. Yang, S. Ong, K. Foong, A robust global and local mixture distance based non-rigid point set registration, *Pattern Recognit.* 48 (2015) 156–173.
- [30] G. Wang, Z. Wang, Y. Chen, W. Zhao, A robust non-rigid point set registration method based on asymmetric gaussian representation, *Comput. Vis. Imag. Underst.* 141 (2015) 67–80.
- [31] T. Kato, S. Omachi, H. Aso, Asymmetric Gaussian and its application to pattern recognition, in: *Structural, Syntactic, and Statistical Pattern Recognition*, Springer, 2002, pp. 405–413.
- [32] R. Jonker, A. Volgenant, A shortest augmenting path algorithm for dense and sparse linear assignment problems, *Computing* 38 (1987) 325–340.
- [33] N. Aronszajn, Theory of reproducing kernels, *Trans. Am. Math. Soc.* 68 (1950) 337–404.
- [34] Y. Zheng, D. Doermann, Robust point matching for nonrigid shapes by preserving local neighborhood structures, *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (2006) 643–649.
- [35] B. Schölkopf, R. Herbrich, A. Smola, A generalized representer theorem, in: *International Conference on Computational Learning Theory*, Springer, 2001, pp. 416–426.
- [36] M. Holden, A review of geometric transformations for nonrigid body registration, *IEEE Trans. Med. Imaging* 27 (2008) 111–128.
- [37] I. Markovsky, *Low Rank Approximation*, Springer, Verlag London, 2012.
- [38] R. Rifkin, G. Yeo, T. Poggio, Regularized least-squares classification, 190, IOS Press, Amsterdam, The Netherlands, 2003.
- [39] J. Ma, J. Zhao, J. Tian, X. Bai, Z. Tu, Regularized vector field learning with sparse approximation for mismatch removal, *Pattern Recognit.* 46 (2013) 3519–3532.
- [40] F. Zhou, F. De la Torre, Factorized graph matching, in: *IEEE Conference on Computer Vision Pattern Recognition*, 2012, pp. 127–134.
- [41] T. Caetano, L. Cheng, Q. Le, A. Smola, Learning graph matching, *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (2009) 1048–1058.
- [42] M. Leordeanu, R. Sukthankar, M. Hebert, Unsupervised learning for graph matching, *Int. J. Comput. Vis.* 96 (2012) 28–45.
- [43] Y. Zheng, D. Doermann, Robust point matching for nonrigid shapes by preserving local neighborhood structures., *IEEE Trans. Pattern Anal. Mach. Intell.* 28 (4) (2006) 643–649.
- [44] J. Ma, J. Zhao, A. Yuille, Non-rigid point set registration by preserving global and local structures, *IEEE Trans. Imag. Process.* 25 (2016) 53–64.
- [45] L. Peng, G. Li, M. Xiao, L. Xie, Robust CPD algorithm for non-rigid point set registration based on structure information., *PLoS ONE* 11 (2) (2016) e0148483.
- [46] D. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2004) 91–110.
- [47] G. Wang, Z. Wang, Y. Chen, W. Zhao, Robust point matching method for multimodal retinal image registration, *Biomed. Signal Process Control* 19 (2015) 68–76.

**Su Zhang** received the B.S. degree from Xiamen University, China in 2015. He is currently working toward the M.S. degrees in the School of Information Science and Technology, Yunnan Normal University. His current research interests include image registration and point set registration.

**Kun Yang** received his master degree from University of New SouthWales, Australia in 1998. He is currently a Professor in the School of Information Science and Technology, Yunnan Normal University and the Engineering Research Center of GIS Technology in Western China of Ministry of Education. His current research interests covers geographic information system and remote sensing image processing.

**Yang Yang** received his master degree from Waseda University, Japan in 2007, and Ph.D. degree from National University of Singapore, Singapore in 2013. His research interest covers image registration, remote sensing image processing, medical image processing, geography information system and human masticatory system.

**Luo Yi** received his B.Sc, M.Sc, and Ph.D. degree all from Harbin University of Science and Technology in 2009, 2012 and 2014, respectively. Now he is a lecturer and master supervisor of Yunnan Normal University. His research interests include wireless sensor networks and geography information.

**Ziqian Wei** is now working toward the B.S.degree in the School of Information Science and Technology, Yunnan Normal University. His current research interests include image registration and point set registration.