Robust Non-Rigid Feature Matching for Image Registration Using Geometry Preserving

Hao Zhu [1,*] , **Ke Zou** [1] , **Yongfu Li** [1] , **Ming Cen** [1] **and Lyudmila Mihaylova** [2]

[1] Key Laboratory of Intelligent Air-Ground Cooperative Control for Universities in Chongqing, and Automotive Electronics and Embedded System Engineering Research Center, College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; kezou18@163.com (K.Z.); liyongfu@cqupt.edu.cn (Y.L.); cenming@cqupt.edu.cn (M.C.)

[2] Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK; L.S.Mihaylova@sheffield.ac.uk

* Correspondence: zhuhao@cqupt.edu.cn

Abstract: In this paper, a robust non-rigid feature matching approach for image registration with geometry constraints is proposed. The non-rigid feature matching approach is formulated as a maximum likelihood (ML) estimation problem. The feature points of one image are represented by Gaussian mixture model (GMM) centroids, and are fitted to the feature points of the other image by moving coherently to encode the global structure. To preserve the local geometry of these feature points, two local structure descriptors of the connectivity matrix and Laplacian coordinate are constructed. The expectation maximization (EM) algorithm is applied to solve this ML problem. Experimental results demonstrate that the proposed approach has better performance than current state-of-the-art methods.

Keywords: image registration; non-rigid feature matching; local structure descriptor; Gaussian mixture model

1. Introduction

Image registration is a fundamental task in many fields, such as computer vision, robotics, medical image processing, and remote sensing [1–4]. The main purpose of image registration is to align two or more images of the same scene taken from different viewpoints, at different times, and/or by different sensors.

Many algorithms have been developed for image registration. It can be roughly divided into area-based and feature-based methods. Area-based methods match image intensity values directly. They mainly include the cross-correlation (CC) methods, the Fourier methods, and mutual information (MI) methods. The normalized CC method is a classic in the area-based methods [5]. The similarity of window pairs from two images are computed and the maximum is considered as a correspondence. A method based on wavelet decomposition and correlation is proposed for image registration [6]. The Fourier methods find the Fourier representation of the images and a subpixel phase correlation with Gaussian mixture model (GMM) is used to register the images [7]. The MI method provides a measure of dependence between two images, and a deterministic explanation for MI-based image registration is proposed in [8].

Feature-based methods extract the salient structures, i.e., features, from the images. The extracted features are called control points [9], and the feature-based methods are considered as point set registration. The traditional for feature-based image registration approach uses a two-step strategy. In the first step, the distances of feature points from local descriptors, such as scale-invariant Fourier

transform (SIFT), speeded-up robust features (SURF), or shape context (SC), are computed to obtain putative correspondences. The putative correspondences may contain some true matches, false matches, and outliers. The true matches from the putative correspondences are then distinguished by some geometrical constraint methods. In [10], an improved random sample consensus (RANSAC) algorithm is proposed to get correct matches, and a common visual pattern is detected for feature matching in pairwise images [11]. In [12], the progressive vector field consensus (VFC) is used to establish the feature points correspondences between images.

The feature-based image registration is formulated to estimate the correspondence matrix and the transformation function. This strategy can be categorized as distance-based methods and probability-based methods [13]. Two-step schemes are employed in distance-based methods. The two point sets distance with the computed correspondence is minimized to estimate the transformation. Widely used methods include the iterative closest point (ICP) algorithm [14] and its variants [15]. Using least squares (LS) minimization, the closest point is considered as the correspondence and the transformation is estimated in the ICP algorithm. In [16], the similarity measures for matching and registration are used as thin-plate spline (TPS) functions in the non-rigid case. In [17], a Havrda–Charvat–Tsallis entropy is proposed to register medical images.

For the probability-based methods, feature points are represented as probability density functions. These methods have attracted considerable attention [18–24]. In the CPD method [18], one point set is represented by Gaussian mixture model (GMM) centroids, and the other point set is fitted to the first point set by moving coherently. In [19], the point sets from medical images are represented as a GMM, the Jensen–Shannon (JS) divergence between these Gaussian densities is minimized for image registration. In [20], both point sets are modeled as GMM, the Euclidean distance between these GMMs is minimized to achieve registration. The above-mentioned methods only consider the global structure between two point sets. In [21], a local descriptor, which is called as a local linear embedding (LLE), is proposed to preserve local structure between two point sets. In [22], a locally linear transforming (LLT), which is similar to LLE, is used to match feature points for remote sensing image. In [24], a dual-feature based point set registration with global-local structural preservation is proposed.

In this paper, we propose a novel GMM-based non-rigid point set registration method to perform image feature matching with structure constraints. The feature points of one image are represented by GMM centroids, and the feature points of the other image are considered as data points. The GMM centroids are moved coherently to fit the data points using a non-rigid transformation by maximizing the likelihood, which encodes the global structure of the pairwise image feature points. For a smooth non-rigid transformation, the coherent constraint is applied by regularization of the displacement field. The shape context is used for the membership probabilities initialization. Furthermore, two local structure descriptors are constructed. The first descriptor is inspired by the idea that the local neighbors in the point set could be preserved after the non-rigid transformation. The Laplacian coordinate is used in the second descriptor to keep the size of the neighborhood structure. Therefore, the objective function is composed of a global distance item, non-rigid transformation constraint item, and two local structure constraint items. An expectation-maximization (EM) algorithm is then applied to perform the maximum likelihood (ML) optimization.

The contribution of this paper can be summarized as follows. (1) Two local structure descriptors of the connectivity matrix and Laplacian coordinates are constructed to preserve the local structure of feature points. The connectivity matrix-based constraint is proposed to retain the local neighborhood relationship, and the Laplacian coordinate-based constraint is employed to encode the local neighborhood scale. The proposed method preserves the local structure in terms of the neighborhood relationship and neighborhood scale, and it is more flexible. (2) The likelihood of image registration is formulated and the EM algorithm is proposed for feature-based image registration.

The paper is organized as follows. The problem of pairwise image registration is formulated in Section 2. The unknown parameters of the proposed method are estimated by the EM algorithm

in Section 3. In Section 4, experiments are carried out to validate the performance of the proposed method. Finally, conclusions are given in Section 5.

2. Problem Formulation

Without loss of generality, the extracted feature points from pairwise images are represented as $\mathbf{X} = \{x_i \,|\, x_i \in \Re^D \}_{i=1}^N$ and $\mathbf{Y} = \{y_j \,|\, y_j \in \Re^D \}_{j=1}^M$, respectively, where D denotes the dimension of the feature points. Our goal is to find a suitable transformation and to establish the correct correspondence between $\mathbf{X}$ and $\mathbf{Y}$. We set the non-rigid transformation of point set $\mathbf{X}$ as the GMM centroids and the other point set $\mathbf{Y}$ as the data point generated by the GMM. That is,

$$p(\mathbf{Y}) = \prod_{j=1}^M \sum_{i=1}^N \pi_{ji} \mathcal{N}(y_j \,|\, f(x_i), \sigma^2 \mathbf{I}_D) \tag{1}$$

where $\mathcal{N}$ denotes the Gaussian distribution, f means the non-rigid transformation, σ^2 represents the equal isotropic covariances, and $\mathbf{I}$ is the identity matrix. We introduce an indicator $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_M]$, where $\mathbf{z}_j$ is a $1 \times M$ binary vector with elements z_{ji} for $i = 1, 2, \ldots N$. The z_{ji} satisfy $z_{ji} \in \{0, 1\}$ and $\sum_i z_{ji} = 1$. It means that only one element in vector $\mathbf{z}_j$ is equal to 1 and all the other elements in vector $\mathbf{z}_j$ are equal to 0. We have

$$p(\mathbf{z}_j \,|\, \boldsymbol{\pi}) = \prod_{i=1}^N \pi_{ji}^{z_{ji}} \tag{2}$$

$$p(y_j \,|\, x_i, \sigma^2, \mathbf{z}_j) = \prod_{i=1}^N \mathcal{N}(y_j \,|\, f(x_i), \sigma^2 \mathbf{I}_D)^{z_{ji}} \tag{3}$$

where $\pi = \{\pi_{ji}\}_{i=1:N}^{j=1:M}$. To account for outliers in $\mathbf{Y}$, a uniform component $p(y_j \,|\, N+1) = \frac{1}{M}$ with weight w is added to the mixture model, where w is a weight of the uniform distribution for noise and outlier. Assuming independent and identically distributed data in $\mathbf{Y}$, we have

$$p(y_j) = w\frac{1}{M} + (1 - w)\sum_{i=1}^N \pi_{ji} \mathcal{N}(y_j \,|\, f(x_i), \sigma^2 \mathbf{I}_D) \tag{4}$$

The negative log-likelihood function of Equation (1) can be represented as:

$$\begin{aligned} Q &= \sum_{j=1}^M \sum_{i=1}^N q(z_{ji}) \frac{\|y_j - f(x_i)\|^2}{2\sigma^2} \\ &+ \frac{D}{2}\log(\sigma^2) \sum_{j=1}^N \sum_{i=1}^N q(z_{ji}) \end{aligned} \tag{5}$$

where $q(z_{ji})$ is the posterior probability.

The non-rigid transformation f is chosen as a displacement function [18]

$$f(\mathbf{X}) = \mathbf{X} + \mathbf{GW} \tag{6}$$

where $\mathbf{G}$ denotes a $N \times N$ dimensional Gaussian kernel matrix with element $g_{ij} = \exp(-\frac{1}{2}\left\|\frac{x_i - x_j}{\tau}\right\|^2)$, τ denotes the width of the Gaussian kernel, and $\mathbf{W}$ is a $N \times D$ dimensional weight matrix of the Gaussian kernel. In order to enforce the motion coherence for preserving the global topology, the constraint on the weight matrix $\mathbf{W}$ given as reference is used [18]

$$E_w(\mathbf{W}) = \mathbf{Tr}(\mathbf{W}^T \mathbf{GW}) \tag{7}$$

where **Tr** denotes the trace of matrix and $\mathbf{W}^T$ is the transposition of matrix **W**. The objective function of feature-based image registration can then be written as

$$
\begin{aligned}
Q &= \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \frac{\|y_j - f(x_i)\|^2}{2\sigma^2} \\
&\quad + \frac{D}{2} \log(\sigma^2) \sum_{j=1}^{N} \sum_{i=1}^{N} q(z_{ji}) + \alpha \mathbf{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{W})
\end{aligned}
\tag{8}
$$

where α is the trade-off parameter.

In Equation (1), the membership probability π_{ji} is chosen in advance. Some methods set the membership probability to be equal [18,25–27]. For robustness, the feature of local shape is proposed here for membership probability initialization. In the 2D case, the shape context (SC) [28] is used as feature descriptor and the Hungarian method is used to perform point matching. In the 3D case, the fast point feature histograms (FPFH) [29] and a sample consensus initial alignment method are used for matching.

Then, two local structure descriptors are constructed to preserve the local structure of point set **X**. By computing the Euclidean distance between each point and its neighbors in **X**, the K nearest neighbors of each point in **X** can be obtained. Each point in **X** can be represented as a weighted linear combination of its K nearest neighbors. Let $\mathbf{L} = \{L_{ij}\}_{i=1:N}^{j=1:N}$ be an $N \times N$ weighted matrix. If a point x_j does not belong to the K nearest neighbors of a point x_i, then L_{ij} is set as 0. The matrix **L** can be obtained by minimizing the cost function:

$$
e(\mathbf{L}) = \sum_{i=1}^{N} \left\| x_i - \sum_{j=1}^{N} L_{ij} x_j \right\|^2
\tag{9}
$$

where the sum of each of row of **L** is equal to 1. After the non-rigid transformation, the local structure can be preserved by minimizing the transformed cost function

$$
\begin{aligned}
&E_l(\mathbf{L}) \\
&= \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \left\| f(x_i) - \sum_{j=1}^{N} L_{ij} f(x_j) \right\|^2 \\
&= \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \left\| x_i + \mathbf{G}(i,.)\mathbf{W} - \sum_{j=1}^{N} L_{ij}(x_i + \mathbf{G}(i,.)\mathbf{W}) \right\|^2
\end{aligned}
\tag{10}
$$

where $\mathbf{G}(i,.)$ means the ith row of **G**. The objective function of feature-based image registration in Equation (8) can be expressed as:

$$
\begin{aligned}
&Q_1 \\
&= Q + \beta E_l(\mathbf{L}) \\
&= \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \frac{\|y_j - f(x_i)\|^2}{2\sigma^2} \\
&\quad + \frac{D}{2} \log(\sigma^2) \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) + \alpha \mathbf{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{W}) \\
&\quad + \beta \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \left\| x_i + \mathbf{G}(i,.)\mathbf{W} - \sum_{j=1}^{N} L_{ij}(x_i + \mathbf{G}(i,.)\mathbf{W}) \right\|^2
\end{aligned}
\tag{11}
$$

where β is the trade-off parameter. For the second structure descriptor, the Laplacian coordinate is proposed to preserve the size of neighborhood structure. For a graph $(\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ and $\mathbf{E}$ is the set of vertices and the set of edges in the graph, respectively, the Laplacian coordinate is expressed as

$$J(v_i) = \sum_{(i,j) \in \mathbf{E}} h_{ij}(v_i - v_j) \tag{12}$$

where (i, j) means the edge between vertices v_i and vertices v_j and h_{ij} is the weight coefficient. The graph of point set X is constructed as follows: the vertex set is set as $\mathbf{V} = \{x_i\}_{i=1:N}$ and the edges (x_i, x_j) if and only if $\|x_i - x_j\|^2 < \varepsilon$, where ε is the threshold parameter to construct the neighborhood graph. h_{ij} is chosen as

$$h_{ij} = e^{-\frac{1}{\varepsilon}\|x_i - x_j\|^2} \tag{13}$$

After the non-rigid transformation, the Laplacian regularization can be preserved by minimizing the following term:

$$
\begin{aligned}
E_h(h) &= \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \|J(x_i) - J(f(x_j))\|^2 \\
&= \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \|J(x_i) - J(x_i + \mathbf{G}(i, .)\mathbf{W})\|^2
\end{aligned}
\tag{14}
$$

The objective function of the feature-based image registration in Equation (11) can be updated as:

$$
\begin{aligned}
Q_2 &= Q_1 + \gamma E_h(h) \\
&= \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \frac{\|y_j - f(x_i)\|^2}{2\sigma^2} \\
&\quad + \frac{D}{2} \log(\sigma^2) \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) + \alpha \mathbf{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{W}) \\
&\quad + \beta \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \left\| x_i + \mathbf{G}(i, .)\mathbf{W} - \sum_{j=1}^{N} L_{ij}(x_i + \mathbf{G}(i, .)\mathbf{W}) \right\|^2 \\
&\quad + \gamma \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) \|J(x_i) - J(x_i + \mathbf{G}(i, .)\mathbf{W})\|^2
\end{aligned}
\tag{15}
$$

where γ is the trade-off parameter.

3. EM for the Proposed Method

In order to estimate the $\Theta = [\mathbf{W}, \sigma^2, \mathbf{Z}]$, the EM algorithm is proposed. There are two steps in the EM algorithm.

(1). $E_L(\Theta, \Theta^{(m)}) = Q_2$

(2). $\Theta^{(m+1)} = \max E_L(\Theta, \Theta^{(m)})$

where m refers to the mth iteration. By iterating these two steps, the parameters Θ are determined while the likelihood function can also be increased.

3.1. E-Step

We use $q(z_{ji})$ to denote $p(z_{ji} = 1 \,|\, \mathbf{Y}, \mathbf{X})$, which can be found using the Bayes' theorem

$$q(z_{ji}) = \frac{\pi_{ji} \mathcal{N}(y_j \,|\, f(x_i), \sigma^2 \mathbf{I}_D)}{\sum\limits_i \pi_{ji} \mathcal{N}(y_j \,|\, f(x_i), \sigma^2 \mathbf{I}_D) + \frac{w}{1-w} \times \frac{N}{M}} \tag{16}$$

3.2. M-Step

The $E_L(\Theta, \Theta^{(m)})$ can be rewritten as

$$
\begin{aligned}
& E_L(\Theta, \Theta^{(m)}) \\
&= \tfrac{1}{2\sigma^2} \left\{ \mathbf{Tr}\,(\mathbf{Y}d(\mathbf{P1})\mathbf{Y}) - 2\mathbf{Tr}\left((\mathbf{P}^T\mathbf{Y})^T (\mathbf{X} + \mathbf{GW}) \right) \right. \\
&\quad \left. + \mathbf{Tr}\left((\mathbf{X} + \mathbf{GW})^T d(\mathbf{P}^T\mathbf{1}) (\mathbf{X} + \mathbf{GW}) \right) \right\} \\
&\quad + \tfrac{D}{2} \log(\sigma^2) \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji}) + \alpha \mathbf{Tr}(\mathbf{W}^T\mathbf{GW}) \\
&\quad + \beta \mathbf{Tr}\,(\mathbf{X}^T\mathbf{BX}) + 2\beta \mathbf{Tr}\,(\mathbf{X}^T\mathbf{BGW}) \\
&\quad + \beta \mathbf{Tr}\,(\mathbf{W}^T\mathbf{GBGW}) \\
&\quad + \gamma \mathbf{Tr}\,(\mathbf{W}^T\mathbf{GS}^T d(\mathbf{P}^T\mathbf{1})\mathbf{SGW})
\end{aligned}
\tag{17}
$$

where $d(.)$ indicates diagonal matrix, $\mathbf{P}$ is the $M \times N$ matrix with element $q(z_{ji})$, $\mathbf{B} = (\mathbf{I} - \mathbf{L})^T d(\mathbf{P}^T\mathbf{1}) (\mathbf{I} - \mathbf{L})$, $\mathbf{1}$ is the all-one column vector, $\mathbf{I}$ is the identity matrix, $\mathbf{S}$ denotes the $N \times N$ Laplacian matrix, $\mathbf{S} = \mathbf{C} - \mathbf{H}$, $\mathbf{H}$ is the $N \times N$ adjacency matrix with element h_{ij}, and $\mathbf{C}$ is the diagonal matrix whose i-th entry is the sum of $(h_{ij})_{j=1,2,\dots N}$.

The estimates of σ_k^2 and $\mathbf{W}$ are updated iteratively by taking the corresponding partial derivative of the expected log likelihood. That is,

$$
\begin{aligned}
& \frac{\partial E_L(\Theta, \Theta^{(m)})}{\partial \sigma^2} \\
&= -\tfrac{1}{2\sigma^4} \left\{ \mathbf{Tr}\,(\mathbf{Y}d(\mathbf{P1})\mathbf{Y}) - 2\mathbf{Tr}\left((\mathbf{P}^T\mathbf{Y})^T (\mathbf{X} + \mathbf{GW}) \right) \right. \\
&\quad \left. + \mathbf{Tr}\left((\mathbf{X} + \mathbf{GW})^T d(\mathbf{P}^T\mathbf{1}) (\mathbf{X} + \mathbf{GW}) \right) \right\} \\
&\quad + \frac{D \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji})}{2\sigma^2} \\
&= 0
\end{aligned}
\tag{18}
$$

We have:

$$
\begin{aligned}
\sigma^2 &= \frac{1}{D \sum_{j=1}^{M} \sum_{i=1}^{N} q(z_{ji})} \left\{ \mathbf{Tr}\,(\mathbf{Y}d(\mathbf{P1})\mathbf{Y}) \right. \\
&\quad - 2\mathbf{Tr}\left((\mathbf{P}^T\mathbf{Y})^T (\mathbf{X} + \mathbf{GW}) \right) \\
&\quad \left. + \mathbf{Tr}\left((\mathbf{X} + \mathbf{GW})^T d(\mathbf{P}^T\mathbf{1}) (\mathbf{X} + \mathbf{GW}) \right) \right\}
\end{aligned}
\tag{19}
$$

Similarly,

$$
\begin{aligned}
& \frac{\partial E_L(\Theta, \Theta^{(m)})}{\partial \mathbf{W}} \\
&= \tfrac{1}{2\sigma^2} \left(-2\mathbf{GP}^T\mathbf{Y} + 2\mathbf{G}d(\mathbf{P}^T\mathbf{1})\mathbf{X} + 2\mathbf{G}d(\mathbf{P}^T\mathbf{1})\mathbf{GW} \right) \\
&\quad + 2\mathbf{GW} + 2\beta\mathbf{GBX} + 2\beta\mathbf{GBGW} \\
&\quad + 2\gamma\mathbf{GS}^T d(\mathbf{P}^T\mathbf{1})\mathbf{SGW} \\
&= 0
\end{aligned}
\tag{20}
$$

$\mathbf{W}$ can be obtained by solving the following system

$$
\begin{aligned}
& \left(d(\mathbf{P}^T\mathbf{1})\mathbf{GW} + 2\sigma^2\mathbf{I} + 2\sigma^2\beta\mathbf{BG} + 2\sigma^2\gamma\mathbf{S}^T d(\mathbf{P}^T\mathbf{1})\mathbf{SG} \right) \mathbf{W} \\
&= \mathbf{P}^T\mathbf{Y} - d(\mathbf{P}^T\mathbf{1})\mathbf{X} - 2\sigma^2\beta\mathbf{BX}
\end{aligned}
\tag{21}
$$

The proposed feature-based image registration algorithm is summarized in Algorithm 1.

Algorithm 1 The proposed non-rigid feature-based image registration algorithm

---

Require:

　The feature point $\mathbf{X} = \{x_i \,|\, x_i \in \Re^D\}_{i=1}^{N}$ and $\mathbf{Y} = \{y_j \,|\, y_j \in \Re^D\}_{j=1}^{M}$, parameters w, α, β, and γ.

1: Initialize $\mathbf{W} = 0$, $\mathbf{P} = \mathbf{I}_{N \times N}$
2: Initialize $\mathbf{W} = 0$, $\mathbf{P} = \mathbf{I}_{N \times N}$
3: Search the K nearest neighbors for each point in $\mathbf{X}$.
4: Perform $\mathbf{L}$ by minimizing the Equation (9)
5: Compute $\{h_{ij}\}_{i=1:N}^{j=1:N}$ as Equation (13)
6: **while** converged **do**

7:　　E-step:
8:　　　Extract the local shape to assign the membership probability π_{ji}
9:　　　Update the $q(z_{ji})$ as Equation (16)
10:　　M-step:
11:　　　Update σ^2 as Equation (19)
12:　　　Update $\mathbf{W}$ by solving the linear system as Equation (21)
13: **end while**

Ensure:

　The aligned point set is $f(\mathbf{X}) = \mathbf{X} + \mathbf{GW}$

　The probability of correspondence is given by $\mathbf{P}$

---

4. Performance Validation

In this section, experiments are carried out to evaluate the performance of the proposed method. The state-of-the-art algorithms ICP method http://www.cvlibs.net/software/libicp/ [14,30], TPS-RPM method https://www.cise.ufl.edu/~anand/publications.html [16], CPD method https://sites.google.com/site/myronenko/research/cpd [18], and CPD-GL method https://sites.google.com/site/jiayima2013/ [31] are used for comparison. The criteria for performance evaluation is chosen as

$$error = \sqrt{\frac{1}{|S|} \sum_{(n,m) \in S} \|f(x_n) - y_m\|^2} \tag{22}$$

where S is the true matches between the two point sets. The *error* refers to the registration error.

4.1. Parameter Settings

In the proposed method, there are seven parameters: w, τ, K, ε, α, β, and γ. Parameter w is a weight of the uniform distribution for noise and outlier. Parameter τ is the width of the Gaussian kernel for non-rigid transformation function. Parameter K is the number of neighbors used to perform the matrix $\mathbf{L}$ in Equation (9) for the first local structure descriptor. Parameter ε is a threshold used to construct the neighborhood graph for the second local structure descriptor. Parameters α, β, and γ are three trade-off parameters for global and local constraints. The other four parameters w, τ, K, ε are selected as [31]. The model selection is shown in Figure 1. In this experiment, the fish dataset is performed and the registration error *error* is chosen as the metric. It is observed that the proposed method has almost the same performance when $\alpha \in [9, 11]$, $\beta \in [200, 400]$, and $\gamma \in [20, 40]$. Therefore, we set $w = 0.1$, $\tau = 2$, $K = 5$, $\varepsilon = 0.05$, $\alpha = 10$, $\beta = 340$, and $\gamma = 24$.
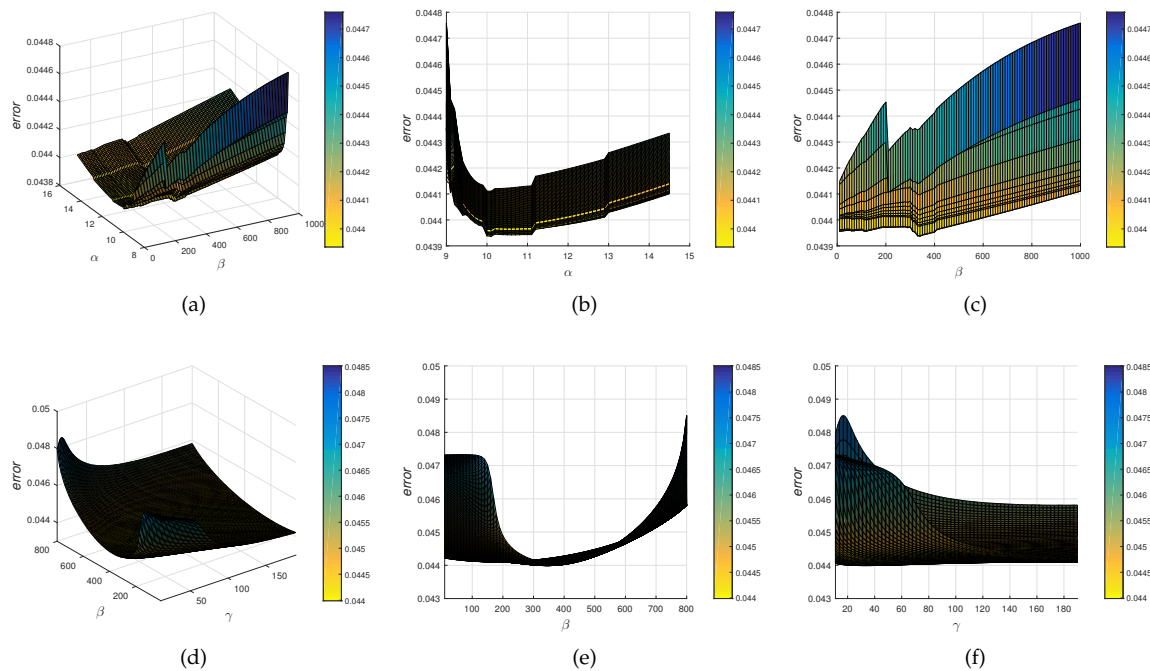
Figure 1. Model selection of the regularization parameters α, β, and γ for structure constraints. (**a**) model selection of the regularization parameters α and β, (**b**) and (**c**) are the different perspectives of (**a**), (**d**) model selection of the regularization parameters β and γ, (**e**) and (**f**) are different perspectives of (**d**).

4.2. Synthesized Point Set Registration

This test is based on the synthesized dataset, which is constructed by Chui and Rangarajan [16]. It consists of two point sets: Chinese character and fish shape. For each model, we designed five sets of data to measure the robustness of registration algorithms with respect to different degrees of deformation and noise. In the deformation test, we use Gaussian radial basis functions to generate deformations, where the coefficients are sampled from a Gaussian distribution with zero mean and standard deviation ranging from 0.02 to 0.08. For the noise test, the standard deviation of the Gaussian noise added to the the original data ratio ranges from 0.01 to 0.05. In each test, one of the above distortions is applied to the model set to create an observed data set, and 100 samples are generated for each degradation level. Comparisons on the fish and Chinese data with varying noise degradation are shown in Figure 2. It is observed that the proposed method with $\beta = 0$ and $\gamma = 0$ has almost the same performance as the CPD-GL method, and the proposed method with $\beta = 340$ and $\gamma = 0$ can register the point set under different degradation slightly better than CPD-GL, whereas the proposed method with $\beta = 340$ and $\gamma = 24$ has the best performance. The metrics under different noise and deformation in fish and Chinese dataset are shown in Figure 3. Apparently, the proposed method with $\beta = 340$ and $\gamma = 24$ has the best performance among all. Furthermore, in order to show the smoothness of the transformation using the proposed method, the experiment of the standard deviation 0.03 for the Gaussian noise added to the Chinese data is performed and 100 samples are generated. The performances of smoothness of the transformation by CPD, CPD-GL, and the proposed method are given in Figure 4, where NP and SD denote the points in the Chinese dataset and the number of samples, respectively, $|f(\mathbf{X})|$ is defined as $|f(\mathbf{X})| = \sqrt{f_x(\mathbf{X})^2 + f_y(\mathbf{X})^2}$, $f_x(\mathbf{X})$ and $f_y(\mathbf{X})$ denote the 1st and 2nd dimensions of $f(\mathbf{X})$, respectively. As can be observed, the proposed method achieves smoother performance than other methods.
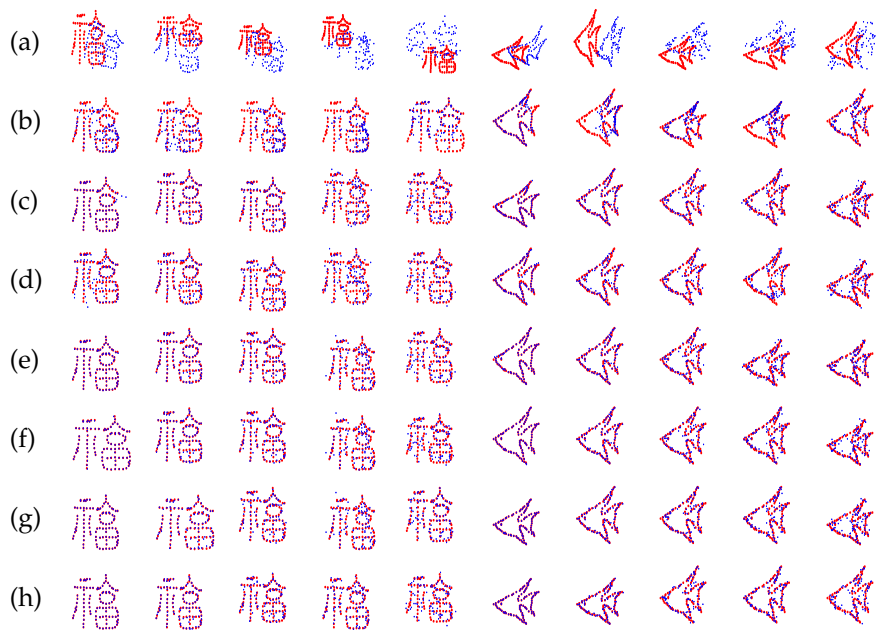
Figure 2. Registration results of point set registration algorithms on the fish and Chinese data with vary noise degradation. (**a**) The data with different gradation level; (**b**) iterative closest point (ICP); (**c**) thin-plate spline (TPS)-RPM; (**d**) CPD; (**e**) CPD-GL; (**f**) the proposed method with $\beta = 0$ and $\gamma = 0$; (**g**) the proposed method with $\beta = 340$ and $\gamma = 0$; (**h**) the proposed method with $\beta = 340$ and $\gamma = 24$.
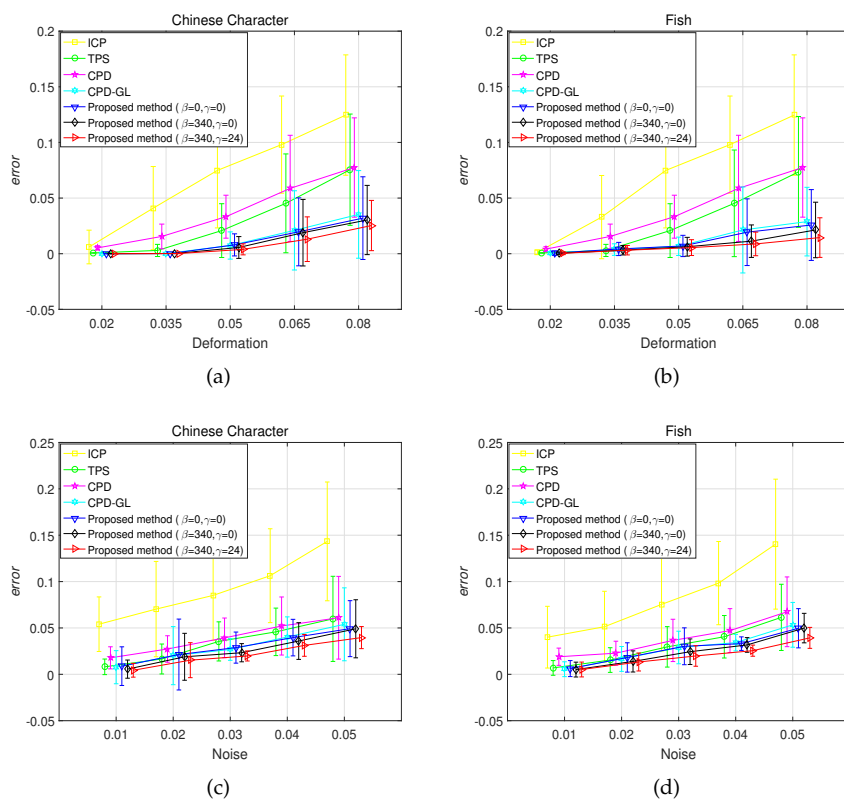

Figure 3. Registration error on Chinese character and fish shape under vary deformation and noise degradation. Sub-figures (**a**) and (**b**) show the registration error on Chinese character and fish shape under vary deformation, respectively, Sub-figures (**c**), and (**d**) show the registration error on Chinese character and fish shape under vary noise degradation, respectively.
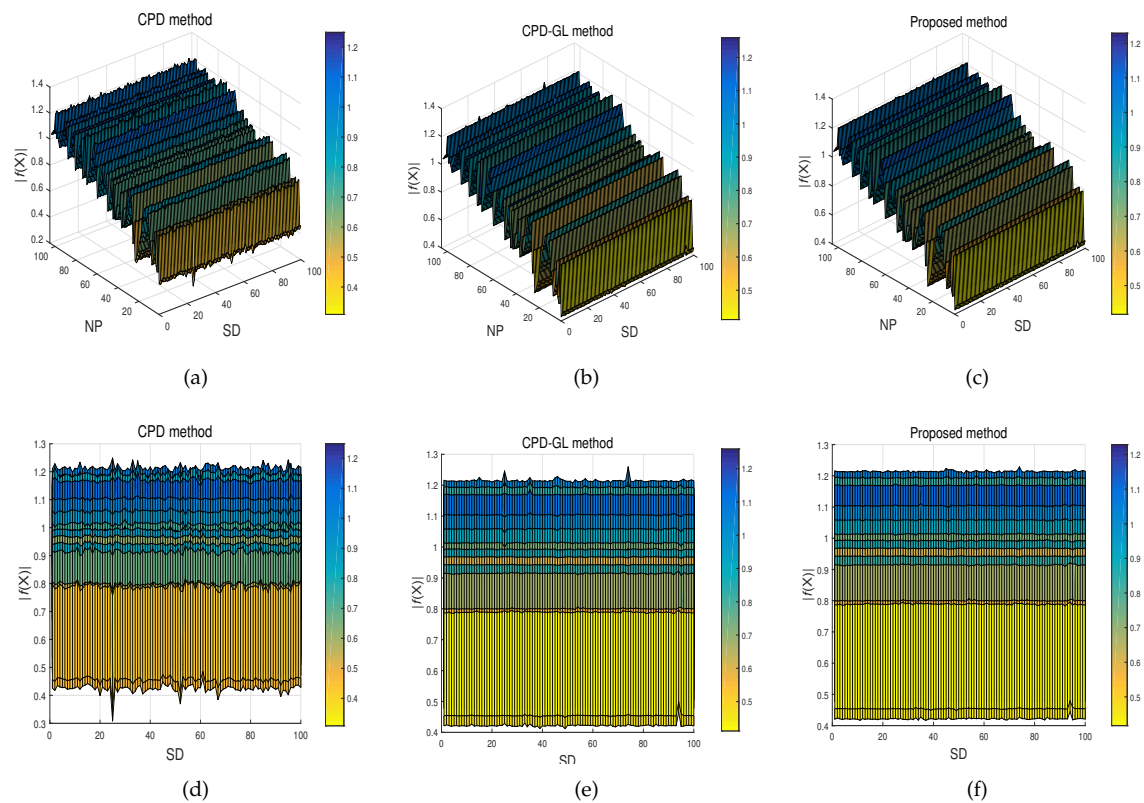
Figure 4. The performances of smoothness of the transformation by CPD, CPD-GL, and the proposed methods. Sub-figures (**a**), (**b**), and (**c**) show the performances of smoothness of the transformation using CPD, CPD-GL, and the proposed methods, respectively, while (**d**), (**e**), and (**f**) depict the same figures in another perspective.

4.3. IMM Hand Landmark Registration

A hand landmark experiment is used to evaluate the performance of the proposed method. The benchmark IMM Hand Database http://www2.imm.dtu.dk/pubdb/views/publication_details. php?id=403 is used. It consists of 40 images of 4 groups of the different human left hand. Images were acquired with resolution 800×600 JPG format in December 2001. Each group has 10 samples with different poses. The hand structures in the dataset have 56 landmarks with the ground truth correspondences. In this experiment, we use 4 groups for evaluation. For each group, the first hand shape is used as the model point set, another 9 hand shapes are considered as data point sets. Therefore, the ground truth correspondences of nine hand landmark pairs for each group are constructed. The proposed method is used to align these face landmark pairs. Comparisons are given in Figure 5. It is observed that the proposed method with $\beta = 340$ and $\gamma = 24$ outperforms than the other methods. Then, comparing the registration error are given in Figure 6, which the performance of the proposed method with $\beta = 340$ and $\gamma = 24$ has the best performance.
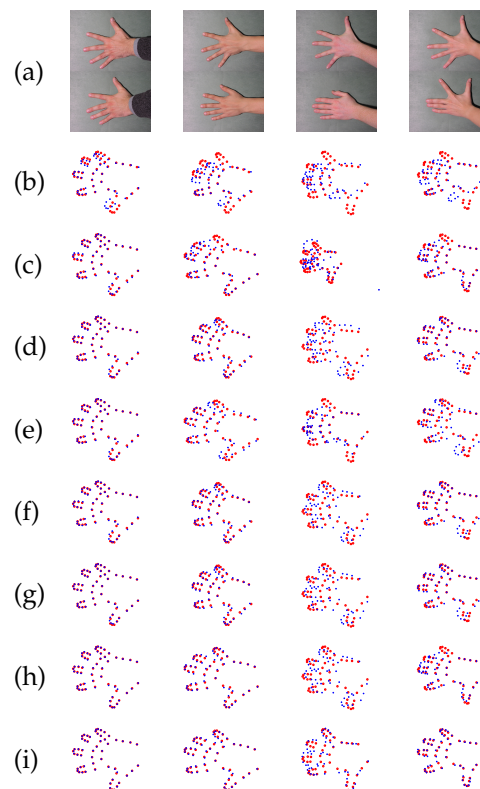
Figure 5. Four examples of IMM hand landmark registration results for each group. (**a**) four hand landmark images (**b**) target hand shape for each group (**c**) ICP (**d**) TPS-RPM (**e**) CPD (**f**) CPD-GL (**g**) the proposed method with $\beta = 0$ and $\gamma = 0$ (**h**) the proposed method with $\beta = 340$ and $\gamma = 0$ (**i**) the proposed method with $\beta = 340$ and $\gamma = 24$.
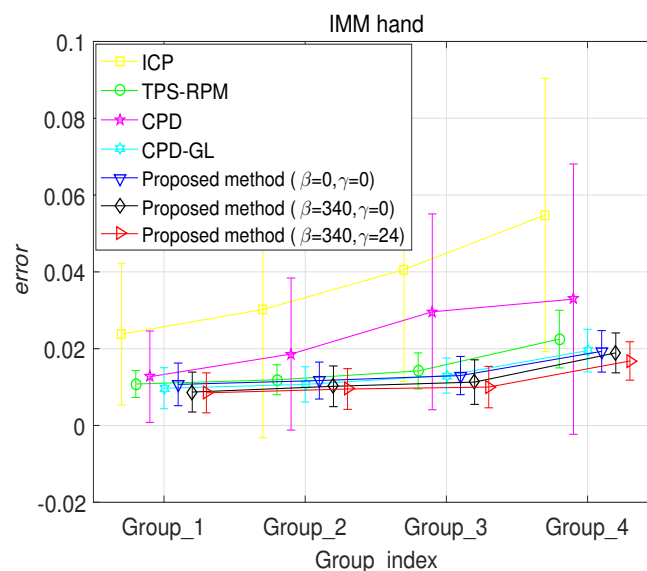

Figure 6. Registration error on IMM hand landmark.

With respect to synthesized point set and IMM hand landmark, CPD-GL outperforms ICP, CPD, and TPS, whereas the proposed method with $\beta = 0$ and $\gamma = 0$ has the almost same performance with CPD-GL method. Moreover, the proposed method with $\beta = 340$ and $\gamma = 0$ has slightly better performance than the CPD-GL method, and the proposed method with $\beta = 340$ and $\gamma = 24$ has the best performance. Then, the computational complexity of the proposed method is analyzed.

The computational complexity for searching the K nearest neighbors for each local track in $\mathbf{X}$ needs $O((K+N)\log N)$ by using the $k-d$ tree, the computational complexity for performing the matrix $\mathbf{L}$ needs $O(K^3 N)$, the EM algorithm needs almost $O(N^3)$. Therefore, the total computational complexity of the proposed algorithm is $O(N^3)$. All algorithms are implemented in MATLAB, and the tests are performed on a Core i7-4790 3.6GHz with 8GB RAM. The run-times of these algorithms are given in Table 1, which illustrates that the proposed method has almost the same computation time with CPD-GL method.

Table 1. Runtime of these point set registration algorithms on different datasets.

Method	Fish	Chinese	IMM Hand
ICP	0.4896s	0.3343s	0.1283s
TPS	2.3457s	1.6755s	0.6231s
CPD	0.1853	0.1278s	0.0954s
CPD-GL	2.2667s	1.4249s	0.3449s
Proposed method	2.4665s	1.7751s	0.4624s

4.4. 4DCT_75 Dataset

The 4DCT_75 dataset of the benchmark data from the DIR-lab http://www.dir-lab.com/index.html is used. The 4DCT_75 data contains ten ("case1" to "case10") thoracic 4D CT images. Each 4D CT images have ten 3D CT images. In particular, we chose "4DCT_75_T00" as a model point set. The other five point sets ("4DCT_75_T10" to "4DCT_75_T50") are set as the scene point sets. Thus, we have 50 pairs of point sets. The example of the registration result of the proposed method are shown in Figure 7, which the proposed method with $\beta = 340$ and $\gamma = 24$ can align these point pairs. Furthermore, the registration errors of CPD method, CPD-GL method, and the proposed method are shown in Figure 8. It is shown that the proposed method with $\beta = 340$ and $\gamma = 24$ has a better performance than the other methods.
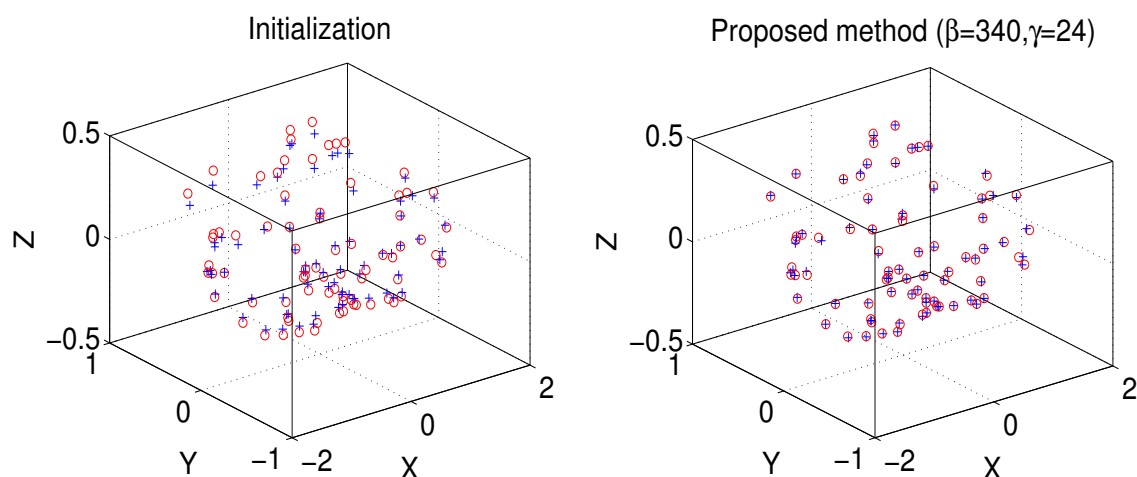

Figure 7. A example of the registration result of the 4DCT_75 data, while the left column denotes the initialization, the right column denotes the registration result using the proposed method with $\beta = 340$ and $\gamma = 24$.
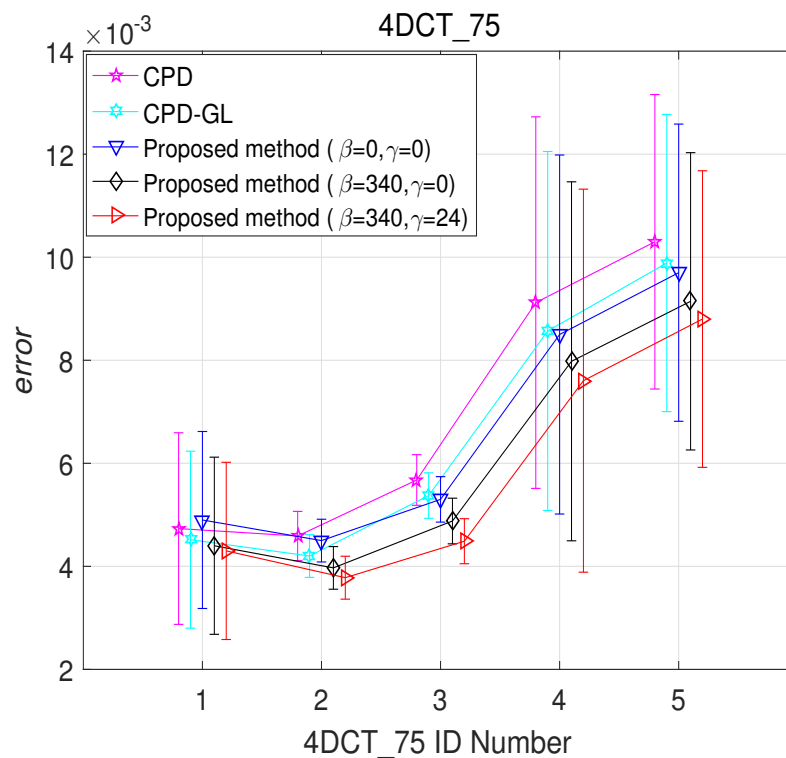
Figure 8. Registration error on "4DCT_75".

4.5. Real Image Feature Matching

In this experiments, the real image pairs are obtained from the data of Jiayi Ma [31], Tuytelaars [32], the Oxford buildings dataset http://www.robots.ox.ac.uk/~vgg/data/oxbuildings, and the Paris dataset http://www.robots.ox.ac.uk/~vgg/data/parisbuildings. The SIFT algorithm is used to extract the feature points of each image, and the putative correspondence is created by nearest neighbor matching. The goal is to find correspondences/matches between two sets of feature points. The performance is evaluated by precision, recall, and F1, where the F1 measure is used to evaluate the balance between recall and precision. The precision, recall, and F1 are defined by: Precision $= \frac{N_L}{N_C}$, Recall $= \frac{N_L}{N_P}$, and F1 $= \frac{2 \cdot \text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}}$, where N_L, N_C, and N_P denote the preserved inlier number, the preserved correspondence number, and the inlier number contained in the putative correspondences, respectively. The ground truth is established in advance manually. The results of matching these image pairs are illustrated in Figure 9. It is observed that the proposed method can distinguish inliers from the outliers. Furthermore, the performance of CPD, CPD-GL and the proposed method in different scenarios is shown in Figure 10. It is observed that the proposed method achieves slightly better performance in precision, recall and F1 metrics compared to other methods. Furthermore, the registration error *error* in pixels is proposed to evaluate the performance of these algorithms. Figure 10d demonstrates that the proposed method performs better than the CPD method and CPD-GL method. Based on the above evaluation results, the proposed method has better performance than other methods.
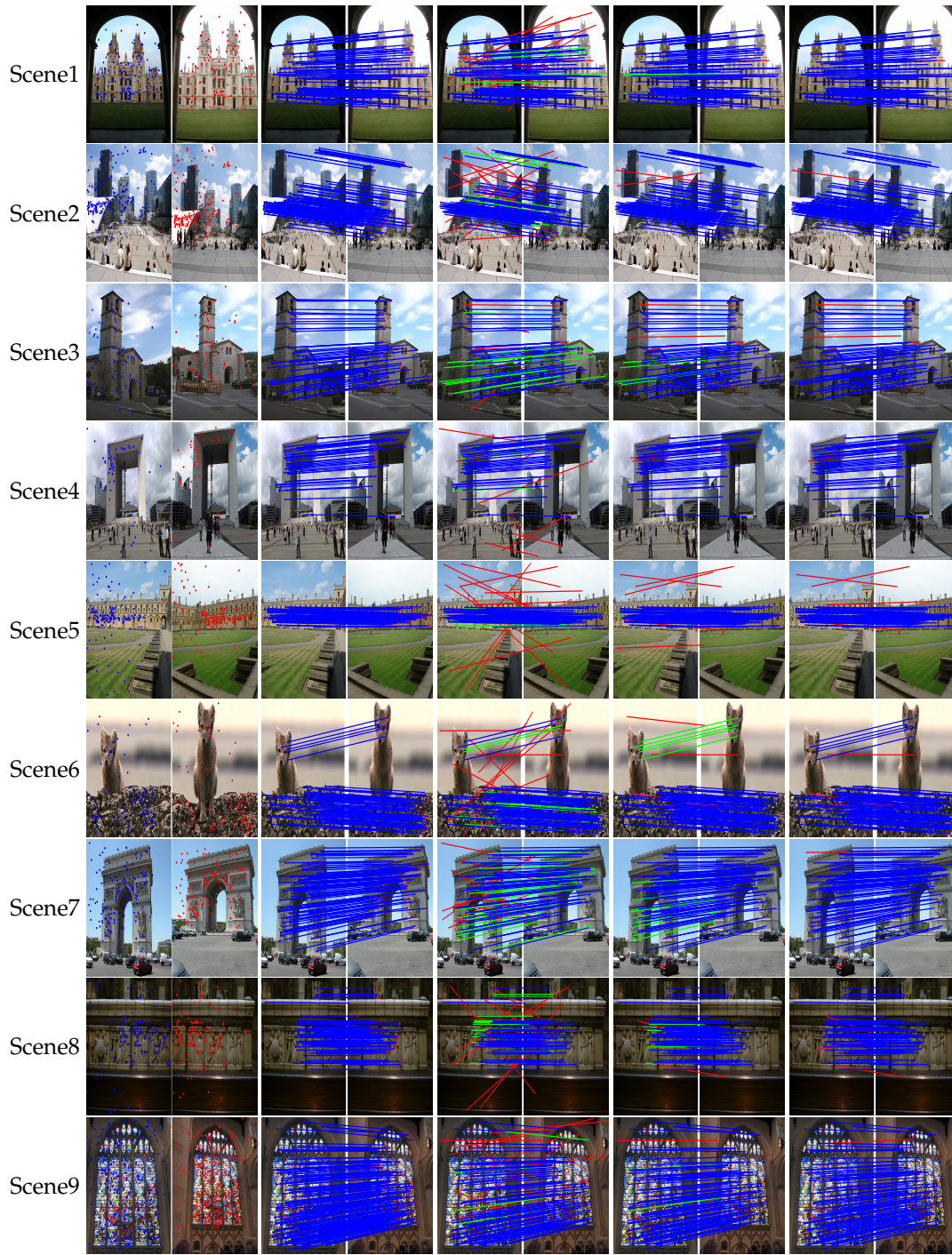
(**a**) Initialization   (**b**) Groud truth   (**c**) CPD   (**d**) CPD-GL   (**e**) Proposed method

Figure 9. Different scenes on real image data. (**a**) different image pairs extracted feature points from scale-invariant Fourier transform (SIFT) algorithm (**b**) manual real correspondence, with results from (**c**) CPD method, (**d**) CPD-GL method, (**e**) proposed method. Each line segment corresponds to the position of the corresponding feature point in the two images (blue = true positive, green = false negative, and red = false positive). For visibility, 100 pairs are selected for some pair of images as presented, and the true negatives are not shown.

(a) Precision

(b) Recall

(c) F1

(d) Registration error (in pixels)
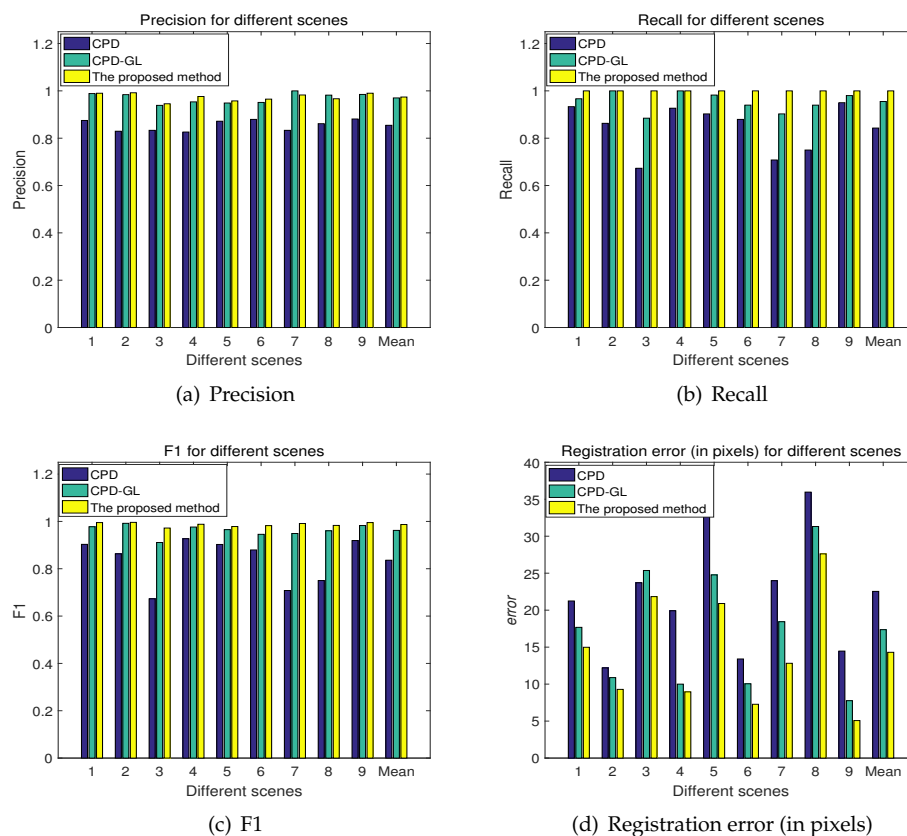
Figure 10. The results of different evaluation metrics of the CPD method, CPD-GL method, and the proposed method in different scenes. (**a**) Precision (**b**) Recall (**c**) F1 (**d**) Registration error (in pixels).

5. Discussion

In this paper, the proposed method is compared with the CPD and CPD-GL methods. The CPD registration method only preserves the global structure and ignores the local structure of point set. Although the CPD-GL algorithm preserves the global and local structure, it only introduces one local structure descriptor. In the proposed method, two local structure descriptors are constructed to preserve the local structure of feature points. The first one is the connectivity matrix-based constraint, which retains the local neighborhood relationship, while the other is the Laplacian coordinate-based constraint, which encodes the local neighborhood scale. Therefore, the proposed method preserves the local structure in terms of the neighborhood relationship and neighborhood scale, and it is more flexible. To illustrate the advantages of this method, three experiments were performed based on the proposed method and compared with the state-of-the-art registration algorithms. In the first experiment, for the 2D point set registration, the experimental results show that the proposed algorithm outperforms other algorithms and has almost the same computation time with CPD-GL algorithm. In the second experiment, registration error is employed to estimate the proposed algorithm on the 3D data, and the proposed algorithm outperforms other algorithms. Furthermore, the proposed algorithm is tested on several image pairs under many different evaluation metrics. The performances of these registration algorithms are evaluated by precision, recall, F1, and registration error. From the above experimental results, the proposed method outperforms the other methods. In particular, it is better to use the proposed method which provides stable and accurate transformation estimation for handling complicated non-rigid deformation. When deformation is not significant, the CPD, CPD-GL, and the proposed method obtain almost similar results, and the CPD has lower computation complexity.

There is room for further development of the proposed framework, especially proposing a fast algorithm and developing more evaluation metrics for point set registration such as smoothness of the transformation. It is interesting to investigate some real-life scenarios, where point set correspondences are weaker such as clouds of points with different amounts of points, missing correspondences or points. Furthermore, the proposed point set registration framework can be applied to perform track-to-track association with sensor bias in distributed multi-target tracking [33].

6. Conclusions

In this paper, a GMM-based probabilistic method for image feature matching is proposed. The feature points of one image are represented by GMM centroids, and the feature points of the other image are fitted to the first image feature points by maximizing the likelihood. To preserve the local structure of these feature points, two local structure descriptors are constructed. The EM algorithm is used to perform image feature matching. Computer simulations confirm the effectiveness of the proposed method in registering different types of images compared to the state-of-the-art registration algorithms.

Author Contributions: Conceptualization, H.Z.; Software, K.Z.; Supervision, L.M.; Writing—original draft, H.Z.; Writing—review & editing, Y.L. and M.C.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhu, H.; Li, Y.; Yu, J.; Leung, H.; Li, Y. Ensemble Registration of Multisensor Images by a Variational Bayesian Approach. *IEEE Sens. J.* **2014**, *14*, 2698–2705. [CrossRef]
2. Li, Y.; He, Z.; Zhu, H.; Zhang, W.; Wu, Y. Jointly registering and fusing images from multiple sensors. *Inf. Fusion* **2016**, *27*, 85–94. [CrossRef]
3. Zhu, H.; Yuen, K.V.; Mihaylova, L.; Leung, H. Overview of Environment Perception for Intelligent Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 2584–2601. [CrossRef]
4. Li, Y.; Tang, C.; Peeta, S.; Wang, Y. Nonlinear Consensus-Based Connected Vehicle Platoon Control Incorporating Car-Following Interactions and Heterogeneous Time Delays. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 2209–2219. [CrossRef]
5. Gonzalez, R.; Woods, R.E. *Digital Image Processing*; 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2002.
6. Moigne, J.L.; Campbell, W.J.; Cromp, R.F. An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 1849–1864. [CrossRef]
7. Dong, Y.; Long, T.; Jiao, W.; He, G.; Zhang, Z. A Novel Image Registration Method Based on Phase Correlation Using Low-Rank Matrix Factorization With Mixture of Gaussian. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 446–460. [CrossRef]
8. Tagare, H.D.; Rao, M. Why Does Mutual-Information Work for Image Registration? A Deterministic Explanation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1286–1296. [CrossRef]
9. Zitova, B.; Flusser, J. Image registration methods: A survey. *Image Vis. Comput.* **2003**, *21*, 977–1000. [CrossRef]

10. Wu, Y.; Ma, W.; Gong, M.; Su, L.; Jiao, L. A Novel Point-Matching Algorithm Based on Fast Sample Consensus for Image Registration. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 43–47. [CrossRef]

11. Liu, H.; Yan, S. Common visual pattern discovery via spatially coherent correspondences. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 1609–1616.

12. Ma, J.; Ma, Y.; Zhao, J.; Tian, J. Image Feature Matching via Progressive Vector Field Consensus. *IEEE Signal Process. Lett.* **2015**, *22*, 767–771. [CrossRef]

13. Zhu, H.; Guo, B.; Zou, K.; Li, Y.; Yuen, K.V.; Mihaylova, L.; Leung, H. A Review of Point Set Registration: From Pairwise Registration to Groupwise Registration. *Sensors* **2019**, *19*, 1191. [CrossRef]

14. Besl, P.J.; McKay, N.D. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [CrossRef]

15. Yang, J.; Li, H.; Campbell, D.; Jia, Y. Go-ICP: A Globally Optimal Solution to 3D ICP Point-Set Registration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 2241–2254. [CrossRef] [PubMed]

16. Chui, H.; Rangarajan, A. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst.* **2003**, *89*, 114–141. [CrossRef]

17. Tustison, N.J.; Awate, S.P.; Song, G.; Cook, T.S.; Gee, J.C. Point Set Registration Using Havrda-Charvat-Tsallis Entropy Measures. *IEEE Trans. Med. Imaging* **2011**, *30*, 451–460. [CrossRef] [PubMed]

18. Myronenko, A.; Song, X. Point Set Registration: Coherent Point Drift. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 2262–2275. [CrossRef]

19. Wang, F.; Vemuri, B.C.; Rangarajan, A.; Eisenschenk, S.J. Simultaneous Nonrigid Registration of Multiple Point Sets and Atlas Construction. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 2011–2022. [CrossRef]

20. Jian, B.; Vemuri, B.C. Robust Point Set Registration Using Gaussian Mixture Models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 1633–1645. [CrossRef]

21. Ge, S.; Fan, G.; Ding, M. Non-rigid Point Set Registration with Global-Local Topology Preservation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, 23–28 June 2014; pp. 245–251.

22. Ma, J.; Zhou, H.; Zhao, J.; Gao, Y.; Jiang, J.; Tian, J. Robust Feature Matching for Remote Sensing Image Registration via Locally Linear Transforming. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6469–6481. [CrossRef]

23. Bai, L.; Yang, X.; Gao, H. Nonrigid Point Set Registration by Preserving Local Connectivity. *IEEE Trans. Cybern.* **2018**, *48*, 826–835. [CrossRef]

24. Zhang, S.; Yang, K.; Yang, Y.; Luo, Y.; Wei, Z. Non-rigid point set registration using dual-feature finite mixture model and global-local structural preservation. *Pattern Recognit.* **2018**, *80*, 183–195. [CrossRef]

25. Horaud, R.; Forbes, F.; Yguel, M.; Dewaele, G.; Zhang, J. Rigid and articulated point registration with expectation conditional maximization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 587–602. [CrossRef]

26. Nguyen, T.M.; Wu, Q.M.J. Multiple Kernel Point Set Registration. *IEEE Trans. Med. Imaging* **2016**, *35*, 1381–1394. [CrossRef]

27. Zhou, Z.; Zheng, J.; Dai, Y.; Zhou, Z.; Chen, S. Robust Non-Rigid Point Set Registration Using Student's-t Mixture Model. *PLoS ONE* **2014**, *9*, e91381. [CrossRef]

28. Salve, S.G.; Jondhale, K.C. Shape matching and object recognition using shape contexts. In Proceedings of the 2010 3rd International Conference on Computer Science and Information Technology, Chengdu, China, 9–11 July 2010; Volume 9, pp. 471–474.

29. Rusu, R.B.; Blodow, N.; Beetz, M. Fast Point Feature Histograms (FPFH) for 3D registration. In Proceedings of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, 12–17 May 2009; pp. 3212–3217.

30. Zhang, Z. Iterative point matching for registration of free-form curves and surfaces. *Int. J. Comput. Vis.* **1994**, *13*, 119–152. [CrossRef]

31. Ma, J.; Zhao, J.; Yuille, A.L. Non-Rigid Point Set Registration by Preserving Global and Local Structures. *IEEE Trans. Image Process.* **2016**, *25*, 53–64.

32. Tuytelaars, T.; Gool, L.V. Matching widely separated views based on affine invariant regions. *Int. J. Comput. Vis.* **2004**, *59*, 61–85. [CrossRef]
33. Zhu, H.; Wang, M.; Yuen, K.; Leung, H. Track-to-Track Association by Coherent Point Drift. *IEEE Signal Process. Lett.* **2017**, *24*, 643–647. [CrossRef]