ORIGINAL ARTICLE

A robust nonrigid point set registration framework based on global and intrinsic topological constraints

Guiqiang Yang[1] · Rui Li[1,2] · Yujun Liu[1,2] · Ji Wang[1,2]

Abstract

The problem of registering nonrigid point sets, with the aim of estimating the correspondences and learning the transformation between two given sets of points, often arises in computer vision tasks. This paper proposes a novel method for performing nonrigid point set registration on data with various types of degradation, in which the registration problem is formulated as a Gaussian mixture model (GMM)-based density estimation problem. Specifically, two complementary constraints are jointly considered for optimization in a GMM probabilistic framework. The first is a thin-plate spline-based regularization constraint that maintains global spatial motion consistency, and the second is a spectral graph-based regularization constraint that preserves the intrinsic structure of a point set. Moreover, the correspondences and the transformation are alternately optimized using the expectation maximization algorithm to obtain a closed-form solution. We first utilize local descriptors to construct the initial correspondences and then estimate the underlying transformation under the GMM-based framework. Experimental results on contour images and real images show the effectiveness and robustness of the proposed method.

Keywords Nonrigid point set registration · Gaussian mixture model · Thin-plate spline · Graph Laplacian regularization · Expectation maximization

1 Introduction

Point set registration [1–4] is a fundamental topic and plays a significant role in machine learning and data mining, and it frequently arises in science tasks, such as augmented reality [5, 6], image mosaicking [7, 8], pattern recognition [9–11], and image fusion [12]. These feature-based image registration tasks work with point features extracted from image pairs, where such a point feature implies some geometric invariance between the images. Image registration can thereby be reduced to a feature point-based registration problem that involves both estimating reliable correspondences and recovering the optimal transformation parameters.

Based on the type of spatial transformations considered, point set registration problems can be divided into problems of rigid registration and nonrigid registration. Rigid registration usually involves rotation, translation, and scaling, and excellent performance has been achieved for this type of registration (e.g., [13–15]). By contrast, nonrigid registration addresses data with various types of degradation (deformation, noise, occlusion, and outliers). It is more challenging because the correspondences between the two point sets are more difficult to establish, and the underlying nonrigid spatial transformations are complicated. However, nonrigid registration arises in many practical engineering, including criminal detection based on palmprint recognition [16], facial expression recognition [17], visual navigation [18], etc. In this paper, we focus on nonrigid registration problem.

A primary challenge in nonrigid registration is how to construct the initial correspondences from the degraded point sets. Here, reliable correspondences are used to refine the optimal spatial transformation, and vice versa. The initial correspondences are usually established based on only local descriptors, e.g., SIFT [19] or shape context (SC) [20], which are inherently sensitive to degradation and hence will give rise to some uncertain correspondences. In brief, the initial correspondences are contaminated by outliers (incorrect correspondences). The goal of transformation estimation

✉ Rui Li
  lirui@dlut.edu.cn

1 School of Naval Architecture, Dalian University of Technology, Dalian 116024, China

2 State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian 116024, China

is to partition the initial correspondences into inliers (correct correspondences) and outliers in accordance with the data points' similarities. However, the existence of outliers makes the registration problem more difficult because spatial transformation learning is a nonlinear registration problem that cannot be modeled in a simple parametric manner. Thus, additional constraints are needed to preserve the global structure and intrinsic geometric structure of the point sets during registration.

To address the above problems, we propose an innovative method for nonrigid point set registration in which the registration problem is converted into a Gaussian mixture model (GMM)-based probability density estimation problem. This method is specifically designed to find an optimal alignment between two given point sets. More precisely, we first construct the initial correspondences using local descriptors. Then, we develop a flexible and robust registration framework for distinguishing inliers from outliers. The key idea of the framework is to employ two complementary constraints in a unified GMM-based optimization framework. One is a thin-plate spline (TPS)-based regularization constraint, which is used to maintain global smoothness during the estimation of the underlying transformation. The other constraint is an intrinsic structure-based geometrical constraint, which is applied to preserve the intrinsic geometry of the transformed set. A joint iterative optimization process that links these two constraints under a GMM-based framework is achieved by using the expectation maximization (EM) [21] algorithm to automatically identify inliers. Moreover, the TPS approach [22] is used to parameterize the spatial transformation, the spline tool TPS has a close-form solution which can be decomposed into affine and nonaffine subspaces. It can produces a smooth functional mapping in the context of supervised learning and has no free parameters that need manual tuning. To generate a smooth spatial mapping between two given point sets with one-to-one correspondences, we choose the TPS for parameterization.

The remainder of this paper is organized as follows. Section 2 reviews related works and the problem we address. Section 3 describes our method in detail. Section 4 presents experimental results and performance comparisons between our approach and other approaches, followed by a conclusion in Sect. 5.

2 Related work

Feature-based image registration is a classical problem that has attracted considerable attention and has been under development for a long time. As mentioned earlier, a feature point-based approach is typically applied in which a local descriptor is used to construct initial corresponding point pairs, and the underlying transformation is then solved for to remove outliers. Here, several theories will be briefly reviewed, and our contributions will be mentioned afterward.

In local descriptor-based methods, a set of initial feature correspondences is built for the extracted point sets by using local neighborhood information. The commonly used descriptors include SIFT, SC, and FPFH [23], the application of which to construct feature correspondences has already been demonstrated. In addition, Hauagge and Snavely [24] presented a technique for finding correspondences between difficult image pairs that are rich in symmetries. Mikolajczyk and Schmid [25] evaluated the performance of a variety of descriptors for local regions of interest and proposed an extension of the SIFT descriptor. However, these local descriptors consider only local neighborhood information, and the initial correspondences usually are either too sparse or contain unavoidable outliers because of ambiguity.

Feature point-based registration is used to align two point sets together and determine the inliers from the initial correspondences. One of most popular registration methods is the iterative closest point (ICP) method [13], which iteratively assigns correspondences between two surface patches under Euclidean transformation. However, as an iterative method, the ICP method itself requires coarse prealignment, without which the method may tend to become trapped in local minima. More critically, such an adequate set of initial poses is no longer valid for nonlinear registration. To alleviate the issue of local minima, various improvements and variants of the standard ICP method have been presented [26–29]. Although these methods can improve the convergence properties of the standard ICP method, they still do not achieve high robustness. Furthermore, a robust feature point-based registration framework based on modeling the transformation with thin-plate spline robust point matching (TPS-RPM) has been presented [30], in which deterministic annealing and soft-assignment techniques are used to learn the transformation. Although the TPS-RPM method is more robust than the ICP method, it has a higher computational complexity. Under this same framework, Tsin and Kanade [31] proposed another trick for point set registration. The kernel correlation (KC) method, also called multiply linked ICP, considers the proportional correlation of two kernel density functions. A later improved approach presented by Jian and Vemuri [32] is based on the KC method; it attempts to model each point set as a mixture of Gaussians and then estimate the transformation by minimizing the discrepancy between the two Gaussian mixtures. However, it is sensitive to outliers. Based on motion coherence theory (MCT) [33, 34], the coherence point drift (CPD) method [35] was introduced and later improved in [36]. In this method, the registration problem is formulated as a maximum likelihood GMM estimation problem with a regularization term for preserving the global topological structure. However, for the CPD method, it is necessary to estimate the underlying

number of Gaussian components, and this method is still sensitive to outliers and noise.

Recently, several methods have been proposed for the detection of outliers through the embedding of spatial constraints and restriction of the solution space. Yang et al. [37] proposed a flexible method called global and local mixture distance with TPS (GLMDTPS), which treats global and local structural differences as a linear assignment problem. The performance of this method which consider only the geometric structure is limited by the assumption that corresponding points have similar structure. Ma et al. [38] proposed a new approach for vector field learning with application to mismatch removing. The SC descriptor was used as local descriptor to assign the membership probabilities, and the EM method was used to solve the regularized mixture model. Yang et al. [39] proposed using the standrad TPS regularization term as prior knowledge to preserve the global topology. However, this approach ignores the intrinsic geometry of the point sets. Wang et al. [40] presented a spatially constrained Gaussian fields (SCGF) method for nonrigid point set registration. They utilized the inner distance shape context (IDSC) [41] to construct the correspondences and used graph Laplacian regularization to constrain the geometric structure of the input space. However, when facing a large degree of degradation, this method is highly time consuming and can easily become trapped in a local minimum. Ma et al. [42, 43] presented an efficient approach based on manifold regularization for the nonrigid registration of shape patterns. However, their method is not robust when confronted with some degree of noise and outlier degradation. Wang et al. [44] formulated the retinal image registration problem as a probabilistic model, which can be used to learn a Gaussian field estimator to achieve robust estimation. Ge and Fan [45, 46] designed a joint optimization approach that combines the CPD constraint and a local linear embedding (LLE) constraint in a GMM-based framework to cope with highly articulated nonrigid deformations, although this method can utilize local structural features more efficiently, the extension performs poorly on data with inhomogeneous density and it more sensitive to outliers because outliers would seriously affect the descriptions of local structures. Thus, their performance degrades in complicated registration problems.

In addition, Ma et al. [47] introduced a simple method for robust feature registration for remote sensing images, called guided locality preserving matching (GLPM), which utilizes the neighborhood structures of potential inliers between the pair images to be registered. Ma et al. [48] introduced a learnable match classifier called LMR for mismatch removal. The classifier constructs match representation for putative matches based on the consensus of local neighborhood structures. Sedaghat and Mohammadi [49] proposed a robust local feature-based registration approach for high-resolution images that is based on the improved SURF detector and localized graph transformation matching. Ma et al. [50] proposed a robust feature guided GMM for image matching which can handle both rigid and nonrigid transformation robustly. Jiang et al. [51] proposed an approach called RFM-SCAN for feature matching, this paper casts the feature matching as a spatial clustering problem with outliers exploiting machine learning techniques. Lati [8] proposed an image mosaicking method for unmanned aerial vehicle images based on random sampling consensus (RANSAC) and bidirectional approaches with a fuzzy inference system.

In summary, the previously mentioned methods have made notable contributions to the field of point set registration, but few of these approaches can achieve good performance when confronted with some degree of degradation. In contrast to the above methods, our proposed method is more robust to degradation than previous methods. The major contributions of the approach proposed in this paper are listed as follows. (1) Global and intrinsic geometric constraints are jointly encoded in a GMM-based registration framework with the aim of preserving the topological structure of the image information. (2) An alternating optimization strategy is used to update the transformation parameters and efficiently eliminate outliers. (3) When this method is applied to contour images and real-world images, the experimental results demonstrate that the method can successfully handle multiple types of degradation.

3 Method

This section elaborates on our method for nonrigid point set registration, which aims to estimate reliable correspondences and learn the underlying transformation between two given point sets subject to ordering information for the model point set $X = \{x_i\}_{i=1}^N \in \mathbb{R}^{N \times D}$ and the scene point set $Y = \{y_j\}_{j=1}^M \in \mathbb{R}^{M \times D}$, where M and N denote the number of points in each point set, respectively, and D is the dimension of the points. The model point set is typically perturbed by deformation, noise, occlusion, and outliers; the hope is that the inliers will still show the maximum pointwise overlap with the scene set. In the context, we consider the global connectivity and intrinsic geometry between two given point sets and propose an efficient nonrigid registration method that tries to identify the inliers among initial correspondences. Specifically, we incorporate two complementary constraints into a GMM-based probabilistic framework. The first constraint is a TPS-based global motion coherence constraint, and the second constraint is a spectral graph-based constraint for intrinsic manifold structure preservation. We can solve the registration problem by minimizing the objective function deduced from the GMM-based problem formulation and these two topological constraints.

3.1 Correspondence estimation

The SC, as an effective descriptor for a point set, is a measure of the distribution of the relative positions of neighboring points. This distribution is defined as a joint histogram in which each histogram axis represents a parameter in polar coordinates [52]. The SC descriptor can be used to search for matches in both 2D and 3D cases.

Furthermore, the local relationship among the neighboring points is well preserved, thus making the matching problem tractable. Logarithmic distance and polar angle bins are used to capture the coarse location information, and these bins are uniform in log-polar space, which makes this feature descriptor more sensitive to the positions of neighborhood points than to those of points farther away [53]. By applying this strategy to establish point correspondences, we can obtain a set $S = \{x_l, y_l\}_{l=1}^{L}$ of putative correspondences between the model and scene point sets, where $L \leq \min\{M, N\}$ is the number of correspondences.

3.2 GMM-based nonrigid registration framework

The putative correspondence set $S = \{x_l, y_l\}_{l=1}^{L}$ will inevitably include mismatches with some unknown outliers. Let $\tilde{X} = \{x_l\}_{l=1}^{L}$ and $\tilde{Y} = \{y_l\}_{l=1}^{L}$. The task is to align $\tilde{X}$ with $\tilde{Y}$ (e.g., $y_l = T(x_l)$) and identify inliers, where T is the optimal transformation function.

In this paper, we focus on identifying as many inliers as possible while achieving satisfactory precision. To this end, we consider a probabilistic mixture model for nonrigid registration. We suppose that the inliers are subject to a multivariate Gaussian distribution with equal isotropic covariances $\sigma^2 I$ and that the outliers satisfy a uniform distribution $1/a$ with a positive constant a. Then, we associate a latent variable $z_l \in \{0, 1\}$ with the l-th corresponding point pair, where $z_l = 1$ for inliers and $z_l = 0$ for outliers. We use homogeneous coordinates for each point, e.g., $x_l = [x_l|1]_{1 \times 3}$, where 1 denotes a matrix element. Thus, the joint probabilistic model can be written as:

$$p(\tilde{Y}|\tilde{X}, \theta) = \prod_{l=1}^{L} \sum_{z_l} p(y_l, z_l|x_l, \theta)$$
$$= \prod_{l=1}^{L} \left[\frac{1-\omega}{a} + \frac{\omega}{2\pi\sigma^2} e^{-\frac{||y_l - T(x_l)||^2}{2\sigma^2}} \right] \quad (1)$$

where $\theta = \{T, \omega, \sigma^2\}$ is a set of unknown latent variables and $\omega \in [0, 1]$ is the mixing coefficient. In accordance with Bayes' theorem, we consider the maximum a posteriori estimate of θ, i.e., $\theta^* = \arg\max_\theta p(\tilde{Y}|\tilde{X}, \theta)$, which can be determined by minimizing the following energy function:

$$E(\theta) = -p(\tilde{Y}|\tilde{X}, \theta) = -\sum_{l=1}^{L} \ln \sum_{z_l} p(y_l, z_l|x_l, \theta) \quad (2)$$

Following the EM algorithm [21], an alternating iterative strategy is considered to address the latent variables. We can find the complete-data logarithm likelihood function of (2) (E step) as follows:

$$Q(\theta, \theta^{old}) = -\frac{1}{2\sigma^2} \sum_{l=1}^{L} p_l ||y_l - T(x_l)||^2$$
$$- \ln \sigma^2 \sum_{l=1}^{L} p_l + \ln \omega \sum_{l=1}^{L} p_l \quad (3)$$
$$+ \ln(1-\omega) \sum_{l=1}^{L} (1 - p_l)$$

where $p_l = p(z_l = 1|x_l, y_l, \theta^{old})$ is a posterior probability indicating the degree to which (x_l, y_l) is an inlier. Based on Bayes' rule, this probability can be obtained by using the old transformation variables as follows:

$$p_l = \frac{\exp\left(-\frac{||y_l - T(x_l)||^2}{2\sigma^2}\right)}{\exp\left(-\frac{||y_l - T(x_l)||^2}{2\sigma^2}\right) + \frac{2\pi\sigma^2(1-\omega)}{\omega a}} \quad (4)$$

Here, we define a threshold ς for identifying inliers $(p_l > \varsigma)_{l=1}^{L}$ after the EM algorithm converges or reaches some termination condition.

The variables are re-estimated, $\theta^{new} = \arg\max_\theta Q(\theta, \theta^{old})$, on the basis of the current p_l. Taking the derivatives of Eq. (3) with respect to σ^2 and ω and setting them equal to zero yields the following (M step):

$$\sigma^2 = \frac{\text{tr}[(\tilde{Y} - T(\tilde{X}))^T P(\tilde{Y} - T(\tilde{X}))]}{2 \cdot \text{tr}(P)} \quad (5)$$

$$\omega = \frac{\text{tr}(P)}{N} \quad (6)$$

where $\text{tr}(\cdot)$ denotes the trace and $P = diag(p_1, \cdots, p_L)$ and $T(\tilde{X}) = (T(x_1), \cdots, T(x_L))^T$. Clearly, we can consider only the relevance of the transformation T and simplify the function (3) to obtain a weighted least squares error functional:

$$Q(T) = \frac{1}{2\sigma^2} \sum_{l=1}^{L} p_l ||y_l - T(x_l)||^2 \quad (7)$$

Function (7) is similar to the moving least square method [54], which can be used to solve for the rigid-body transformation T. However, for correlates of coherent visual motion such as changes in facial expression and multiview image sequences, for which the feature points are extracted in the presence of degradations, the principle of

as-rigid-as-possible image registration may not work well. The nonrigid registration problem is essentially a regression problem. It may turn out that the problem of directly solving for the transformation T in (7) is ill-posed since no unique solution is available. Regularization techniques, which usually operate in a reproducing kernel Hilbert space (RKHS), are often used to produce reasonable solutions to ill-posed problems.

3.3 Global structure preservation

In order to complete the transformation estimation for the nonrigid registration, we convert the formulation into the nonrigid case by applying regularization theory and the optimal nonrigid transformation T can then be estimated by minimizing a weighted regularized least squares error functional:

$$\mathcal{E}(\mathrm{T}) = \frac{1}{2\sigma^2} \sum_{l=1}^{L} p_l ||y_l - \mathrm{T}(x_l)||^2 + \frac{\gamma}{2}\phi(\mathrm{T}) \tag{8}$$

where $\gamma \geq 0$ is a fixed regularization parameter that controls the trade-off and $\phi(\mathrm{T})$ is the registration term. The transformation T is a vector-valued function that lies in a RKHS. Thus, the functional ϕ has the form $\phi(\mathrm{T}) = ||\mathrm{T}||_{\mathcal{H}}^2$, where $\mathcal{H}$ is the RKHS and $|| \cdot ||_{\mathcal{H}}$ denotes the norm $\mathcal{H}$.

To complete the least squares estimation of the transformation, we consider a nonrigid transformation function of a specific form, namely a TPS to generate a smooth mapping fitting. A TPS is a general purpose spline tool for representing a flexible nonrigid transformation, which has the advantages of free parameters and a closed-form solution. We model the transformation T using the TPS approach, in which the resulting transformation can be decomposed into a global affine transformation and a local nonaffine warping component:

$$\mathrm{T}(x_l) = x_l \cdot \mathrm{H} + k(x_l, x) \cdot \mathcal{C} \tag{9}$$

where H is defined as a 3×3 global affine matrix, $\mathcal{C}$ is a $L \times 3$ nonaffine coefficient matrix, and the radial basis function $k(x_l, x)$ is a $1 \times L$ vector defined by the TPS kernel for each point, i.e., $k(x_l, x) = ||x_l - x||^2 \log ||x_l - x||$. The TPS kernel matrix $\mathcal{K}_{L \times L}$ is defined, with each element expressed as $\mathcal{K}_{lp} = k(x_l, x_p) = ||x_l - x_p||^2 \log ||x_l - x_p||$, where x_p is a point in $\tilde{\mathrm{X}}$.

In motion perception, the concept of motion coherence [33, 34] is applied to intuitively interpret points close to one another as tending to move coherently in space. Motion coherence is a particular way of imposing smoothness on a spatial transformation. We can regularize the local nonaffine component $k(x_l, x) \cdot \mathcal{C}$ to enhance the motion coherence and preserve the global structure; then, the regularization term has the form:

$$||\mathrm{T}||_{\mathcal{H}}^2 = \left\langle \sum_{l=1}^{L} k(x_l, x)\mathcal{C}, \sum_{l=1}^{L} k(x_l, x)\mathcal{C} \right\rangle_{\mathcal{H}}$$
$$= \sum_{l=1}^{L} \langle k(x_l, x_l)\mathcal{C}, \mathcal{C} \rangle = \mathrm{tr}(\mathcal{C}^{\mathrm{T}}\mathcal{K}\mathcal{C}) \tag{10}$$

Substituting Eq. (10) into Eq. (8), we can obtain a TPS energy function as follows:

$$\mathcal{E}(\mathrm{H}, \mathcal{C}) = \frac{1}{2\sigma^2}||\mathrm{P}^{1/2}(\tilde{\mathrm{Y}} - \tilde{\mathrm{X}}\mathrm{H} - \mathcal{K}\mathcal{C})||^2 + \frac{\gamma}{2}\mathrm{tr}(\mathcal{C}^{\mathrm{T}}\mathcal{K}\mathcal{C}) \tag{11}$$

where the second term is the standard TPS regularization term, which is used to maintain the overall spatial relationship of the point sets during registration. It is solely dependent on the nonaffine subspace.

The transformation learning method applied here is a dimension reduction technique on the complete subspace, which aims to produce a smooth function T. However, it fails to discover the underlying intrinsic geometry of the marginal space X, which is essential in real applications.

3.4 Intrinsic structure preservation

The main aim of nonrigid registration is to estimate a transformation that maintains the global connectivity and associativity of the entire point set and preserves the intrinsic manifold structure during the transformation. In most nonrigid registration tasks, the success of registration learning for some number of the total set of point pairs (i.e., $L \leq N$) can be plausibly achieved through the effective use of the labeled point pairs $\{x_l, y_l\}_{l=1}^L$. However, the remaining $N - L$ unlabeled point pairs may contain additional structure information, which can impose smoothness conditions on possible solutions. To preserve the intrinsic geometry of the data space, leveraging recent advances in spectral graph theory, we attempt to extend the energy function in Eq. (11) by incorporating ensemble graph Laplacian regularization, thereby ensuring that the transformation is smooth with respect to both the RKHS $\mathcal{H}$ and the marginal distribution.

Motivated by the manifold assumption [55], if two feature points x_i and x_j are close in the intrinsic geometry of the marginal distribution on the entire set of transformation data $\mathcal{T}(\mathrm{X}) = (\mathcal{T}(x_1), \cdots, \mathcal{T}(x_N))^{\mathrm{T}}$, then the representations of these two points in the new basis should also be close when the Laplacian regularization term is minimized. The graph Laplacian [56] can be used to discretely approximate the manifold Laplacian $||\mathcal{T}||_{\mathcal{M}}^2$, which measures the smoothness of $\mathcal{T}$ along the data manifold on the model point set X.

Consider a graph that is conventionally defined as $\mathcal{G} = \{\mathcal{T}, r, \mathrm{W}\}$, where $\mathcal{T}$ is a set of vertices that can be written in matrix form as $\mathcal{T} = \mathrm{XH} + \mathrm{KC}$, where $\mathrm{K}_{N \times N}$ is the Gram matrix and $\mathrm{K}_{ij} = k(x_i, x_j)$, $\mathrm{C}_{N \times 3}$ is the matrix of coefficients;

r, which is a constant parameter, is used to compute the weights of the edges of the graph, and W is an edge-weight matrix, with W_{ij} being the weight of the edge connecting vertices $\mathcal{T}(x_i)$ and $\mathcal{T}(x_j)$:

$$W_{ij} = \begin{cases} e^{-\frac{1}{r}||\mathcal{T}(x_i)-\mathcal{T}(x_j)||^2} & if ||\mathcal{T}(x_i) - \mathcal{T}(x_j)||^2 \le r \\ 0 & \text{otherwise} \end{cases} \quad (12)$$
$$\text{s.t.} \quad \mathcal{T}(x_j) \in \mathcal{N}(\mathcal{T}(x_i))$$

where the neighborhood $\mathcal{N}(\mathcal{T}(x_j))$ is the set of vertices connected to vertex $\mathcal{T}(x_i)$ by edges. According to W, the graph Laplacian is a real symmetric matrix, and it has a complete set of orthonormal eigenvectors that can be defined as $A = D - W$, where D is a diagonal matrix with elements $D_{ii} = \sum_{i=1}^{N} W_{ij}$. Thus, the data manifold Laplacian can be preserved by minimizing the following:

$$||\mathcal{T}||_{\mathcal{M}}^2 = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} W_{ij}||\mathcal{T}(x_i) - \mathcal{T}(x_j)||^2 = \text{tr}(\mathcal{T}^T A \mathcal{T}) \quad (13)$$

As seen from Eq. (13), which is defined on the basis of a nonnormalized graph Laplacian that is sensitive to the vertex degree of the graph, it is reasonable to use the normalized graph Laplacian $\tilde{A} = D^{-1/2}AD^{-1/2}$ instead of A in all formulas.

3.5 Global-intrinsic structure preservation

We assume that $\tilde{X}$ and $\tilde{Y}$ are in the putative point set S and correspond to the first L rows of the given point sets X and Y, respectively. Based on the TPS energy function (11) and the given graph Laplacian regularization term, the transformation parameters can be estimated by minimizing the following function:

$$\mathcal{E}(H, C) = \frac{1}{2\sigma^2}||P^{1/2}(\tilde{Y} - \tilde{X}H - JKC)||^2 + \frac{\gamma_1}{2}\text{tr}(C^T J^T \mathcal{K} JC) + \frac{\gamma_2}{2}\text{tr}(\mathcal{T}^T \tilde{A}\mathcal{T}) \quad (14)$$

where $J_{L \times N} = (I_{L \times L}, 0_{L \times (N-L)})$, with I being a unit matrix and the rest of the elements being 0. Solving for the parameter pair H and C by applying an iterative updating method in which we take the derivatives of (14) with respect to H and C and set them equal to zero, which leads to the following linear system:

$$(\tilde{X}^T P\tilde{X} + \gamma_2\sigma^2 X^T \tilde{A}X)H = (\tilde{X}^T P\tilde{Y} - \tilde{X}^T PJKC - \gamma_2\sigma^2 X^T \tilde{A}KC) \quad (15)$$

$$(K^T J^T PJK + \gamma_1\sigma^2 J^T \mathcal{K} J + \gamma_2\sigma^2 K^T \tilde{A}K + \tilde{\varepsilon}I)C = (K^T J^T P\tilde{Y} - K^T J^T P\tilde{X}H - \gamma_2\sigma^2 K^T \tilde{A}XH) \quad (16)$$

where $\tilde{\varepsilon}$I control the stability of the numerical value. Thus, we obtain a stable solution for $\mathcal{T}$ in Eq. (14).

3.6 Fast implementation

The kernel matrix scales poorly for high-dimensional nonrigid registration, which may pose a serious problem because of high computational complexity in time and space. To address this problem, we provide an approximation method as the basis for a fast implementation, with the aim of reducing the computational complexity, based on Ref. [57]. Give a subset $\{\hat{x}_q\}_{q=1}^{Q} (Q \ll L)$ drawn from the putative point set $\tilde{X}$ at random without replacement. Then, we consider replacing $\mathcal{T}$ in Eq. (14) by searching for a solution that allows only the coefficients $\hat{C}$ to be nonzero. With slightly different notation, we seek a solution that has the following approximate solution:

$$\mathcal{T}(\hat{x}_q) = \hat{x}_q \cdot \hat{H} + k(\hat{x}_q, \hat{x}) \cdot \hat{C} \quad (17)$$

where $k(\hat{x}_q, \hat{x})$ is a $1 \times Q$ vector, $\hat{H}$ is a 3×3 matrix and $\hat{C}$ is an $Q \times 3$ matrix. We obtain the optimal coefficients $\hat{H}$ and $\hat{C}$ from the following linear system:

$$(\tilde{X}^T P\tilde{X} + \gamma_2\sigma^2 X^T \tilde{A}X)\hat{H} = (\tilde{X}^T P\tilde{Y} - \tilde{X}^T PJU\hat{C} - \gamma_2\sigma^2 X^T \tilde{A}U\hat{C}) \quad (18)$$

$$(U^T J^T PJU + \gamma_1\sigma^2 G + \gamma_2\sigma^2 U^T \tilde{A}U + \tilde{\varepsilon}I)\hat{C} = (U^T J^T P\tilde{Y} - U^T J^T P\tilde{X}\hat{H} - \gamma_2\sigma^2 U^T \tilde{A}X\hat{H}) \quad (19)$$

where U is the TPS kernel matrix $U \in \mathbb{R}^{N \times Q}$ with elements $U_{iq} = k(x_i, \hat{x}_q)$ and $G \in \mathbb{R}^{Q \times Q}$ with elements $U_{ij} = k(\hat{x}_i, \hat{x}_q)$. Compared to the linear system in Eq. (16), the time complexity of computing the coefficients in Eq. (19) is reduced from $\mathcal{O}(N^3)$ to $\mathcal{O}(MN^2)$. The spatial complexity of our fast implementation can be considered to be $\mathcal{O}(N^2)$ due to the memory requirements for storing the graph Laplacian matrix $\tilde{A}$.

3.7 Implementation details

The registration performance of our method typically depends on the homogeneous coordinate system, in which the coordinates of the points in both sets have zero mean and unit variance.

Parameter settings: There are six main parameters in our method, and all six parameters are set on the basis of experimental evaluations. The regularization parameters include γ_1 and γ_2, which are used to control the effects of the global and intrinsic topological constraints, respectively, on the transformation $\mathcal{T}$; we fix them to values of $\gamma_1 = 5$ and $\gamma_2 = 0.05$. The parameter a of the uniform distribution is set to $a = 10$. For the inlier ratio ω of the Gaussian distribution, we adopt the initial assumption that this parameter has a fixed value of

0.9. The parameter ς, which is used to identify the reliability of the inliers, is fixed to 0.75. The parameter r, which is used to construct the edge-weight matrix, is empirically set to 0.05. We summarize our approach in algorithm 1.

Algorithm 1. The pseudocode of our method

Input: Two point set $X = \{x_i\}_{i=1}^{N}$ and $Y = \{y_j\}_{j=1}^{M}$;

Output: Aligned point set $\{\mathcal{T}(x_i)\}_{i=1}^{N}$.

1 Require: γ_1, γ_2, a, ς, and r;

2 **Repeat**

3   Compute $S = \{(x_l, y_l)\}_{l=1}^{L}$ using the SC descriptor;

4   Construct U and G based on their definitions;

5   Initialization: $P = I_{L \times L}$, $\hat{H} = I_{3 \times 3}$, $\hat{C} = 0_{Q \times 3}$,

    $\omega_{initial} = 0.9$, and σ^2 is determined by Eq. (5);

6 **Repeat**

7   Compute matrix $\tilde{A}$ using Eq. (12);

8   E step: Update P using Eq. (4);

9   M step: Update $\hat{H}$ and $\hat{C}$ using Eqs. (18) and (19);

    Update σ^2 and ω using Eqs. (5) and (6);

10 **Until** a specified termination condition is reached;

    Transformation point set: $\{x_i\}_{i=1}^{N} \leftarrow \{\mathcal{T}(x_i)\}_{i=1}^{N}$

11 **Until** reach the maximum iteration number;

12 **Return** $\{\mathcal{T}(x_i)\}_{i=1}^{N}$ in the last iteration.

---

4 Experiments and analysis

In this section, we evaluate the performance of our proposed method on both contour images and real-world images for comparison with other state-of-the-art methods: GLTP [45, 46], EM-TPS [39], MR-RPM [41, 42], GMMREG [32], and CPD [35]. All tests were run directly in the MATLAB environment on a standard PC with an Intel i5 3.4 GHz CPU.

We consider the registration error to evaluate the performance of the registration methods. The error is defined as the root mean square error, which directly characterizes the registration performance; it is calculated as follows:

$$\text{Error} = \sqrt{\frac{1}{N_{\text{total}}} \sum_{k=1}^{N_{\text{total}}} (y_k - \mathcal{T}(x_k))^2} \tag{20}$$

where N_{total} denotes the total number of true correspondences and y_k is the ground truth correspondence with x_k.

4.1 Overall performance on contour images

The contour images considered here [30] are based on two category shape models: a *fish* and a *Chinese character*. The *fish* and *Chinese character* datasets contain 98 and 105 points, respectively. For each shape model, four groups of point sets with different types of degradation, including deformation, noise, occlusion, and outliers, are designed to test the performance of various registration methods. For the *fish* and *Chinese character* datasets, there are five or six

levels of degradation, and 100 samples are generated for each degradation level.

4.1.1 Nonrigid fish point set registration

The registration results obtained on the *fish* shape model are illustrated in Fig. 1 for the largest degree of degradation in each category, such as deformation level to 0.08, noise level to 0.05, occlusion ration to 0.5, and outlier ratio to 2. As shown in these results, MR-RPM and our method perform well for nonrigid deformation. In the noise experiment, we see that EM-TPS can align the two point sets. However, it fail to generate the correct correspondence. In the occlusion and outlier tests, our method gets superior registration results.

To further explore the registration performance, we consider the error bars of the registration results to evaluate the performance of the registration methods. The error bars represent the means and standard deviations of the registration errors on all 100 samples for each level of the various types of degradation. We summarize the experimental results in Fig. 2. According to Fig. 2a, we observe that our method and MR-RPM perform better due to adopting the intrinsic topological constraints. In Fig. 2b, there is no winner when the model point set is disturbed by the white Gaussian noise. Nevertheless, our method can achieve quite satisfactory results, followed by MR-RPM. Figure 2c and d show the results for the cases of occlusion and outliers. Our method performs bests because of constructing correspondences for contour points under
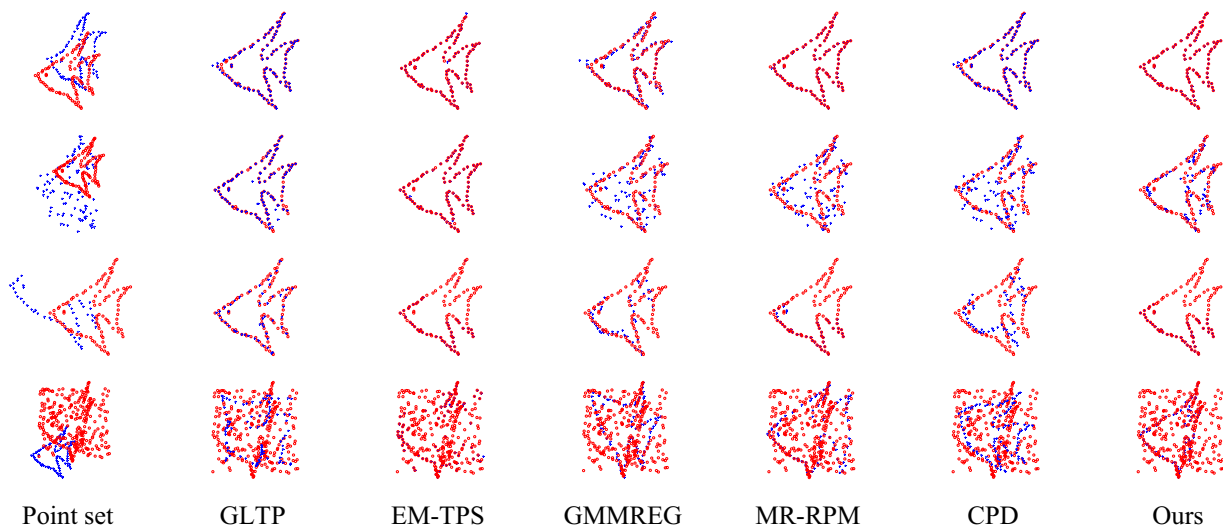

| Point set | GLTP | EM-TPS | GMMREG | MR-RPM | CPD | Ours |

Fig. 1 Point set registration results for *fish* shape samples with the largest degradations, from the top row to the bottom row: deformation (0.08), noise (0.05), occlusion (0.5), and outliers (2). The leftmost column shows the model point set (blue pluses) and the scene point set (red circles). The goal is to align the model point set with the scene point set. In the second column from the left to the rightmost column, the registration of GLTP, EM-TPS, GMMREG, MR-RPM, CPD, and our method, respectively, is presented
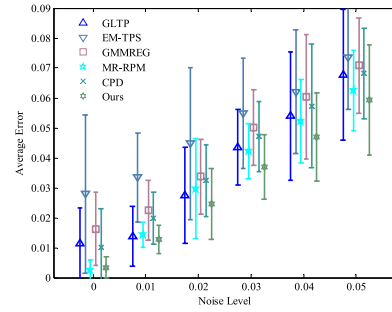
Fig. 2 Comparison of the performance of the registration methods for point sets based on the *fish* shape model with different levels of degradation. **a**, **e**: Deformation test. **b**, **f**: Noise test. **c**, **g**: Occlusion test. **d**, **h**: Outlier test
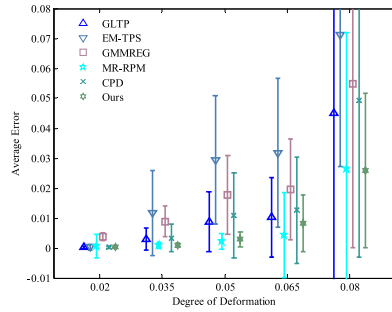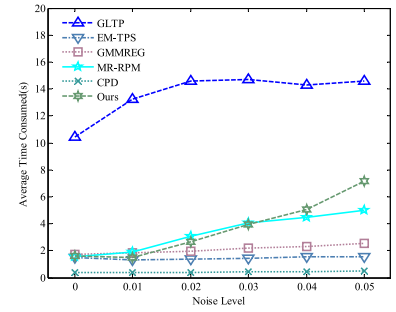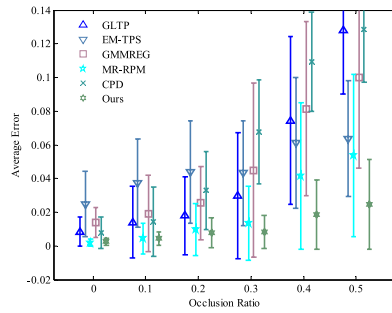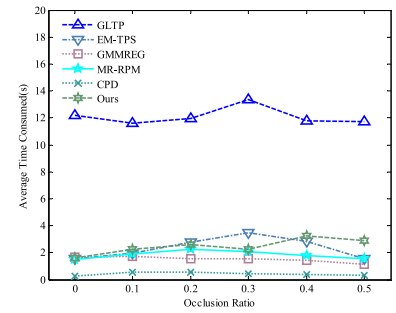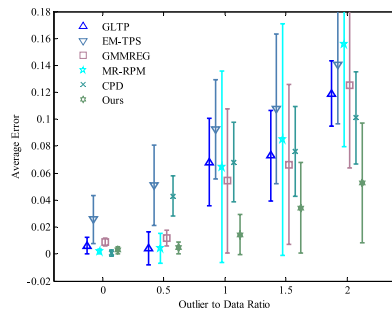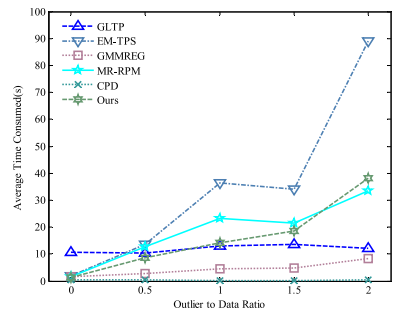

(a)
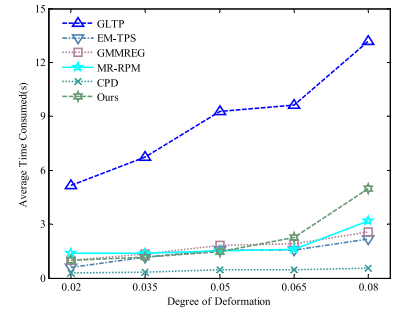
(e)

(b)

(f)

(c)

(g)

(d)

(h)

an ordering constraint and integrating two topologically complementary constraints into a unified optimization framework. Figure 2e–h separately show the average time consumptions of the different registration methods at the different degradation levels. We can see GLTP spend more time under deformation, noise, and occlusion test. Meanwhile, in the outlier test, the more time consumption of EM-TPS, MR-RPM and our proposed method when the outlier level becomes relatively high, such as 1.5 or 2.

In summary, the registration performance of all methods gradually declines as the degree of degradation increases. Our method performs more desirably in general because of integrating global and intrinsic topological constraints into a GMM-based framework and modeling the transformation using the TPS approach. GLTP performs wells in the deformation and noise tests. However, it is sensitive to occlusion and outliers because the LLE constraints are less reliable when the point sets are sparse or nonuniform. The EM-TPS method performs poorly in most cases, and its registration error shows strong fluctuations because it only consider the global distribution of the model point set. The MR-RPM method achieves good performance in deformation and noise tests because of the manifold regularization technique. However, it is vulnerable to occlusion and outliers, because the initial correspondences are likely inaccurate. GMMREG is not robust in most cases, especially when the data point set is badly degraded. The alignment performance of CPD is relatively good in deformation and noise tests and spends

less time than others in all tests, but it is still sensitive to occlusion and outliers due to the impossibility of capturing the true distributions of the point sets.

4.1.2 Nonrigid the Chinese character point set registration

The performance in the *Chinese character* experiment is displayed in Fig. 3 for the highest level of each type of degradation. Our method and MR-RPM perform better than the other four methods in the deformation test. In the noise test results, we observe that the model point set is distorted by Gaussian noise, which disturbs the neighborhood structures in the model point set. It is extremely difficult for all methods to achieve accurate alignment. When there are missing points or outliers in the model point set, we observe that our method gives the best registration results.

Quantitative comparative experiment results on the *Chinese character* dataset under four degradation scenarios are shown in Fig. 4. In Fig. 4a, the average error of all methods gets larger as the level of degradation goes up. The growth rate of the average error is much slower for MR-RPM and our method than for the other methods. In Fig. 4b, all methods except EM-TPS can obtain satisfactory results, our proposed method performs slightly better than the other tested methods with the noise level ranging from 0 to 0.05. Figure 4c and d present the registration results of all methods in the presence of occlusion and outliers, respectively. GLTP, EM-TPS, GMMREG, MR-RPM, and CPD are not robust to
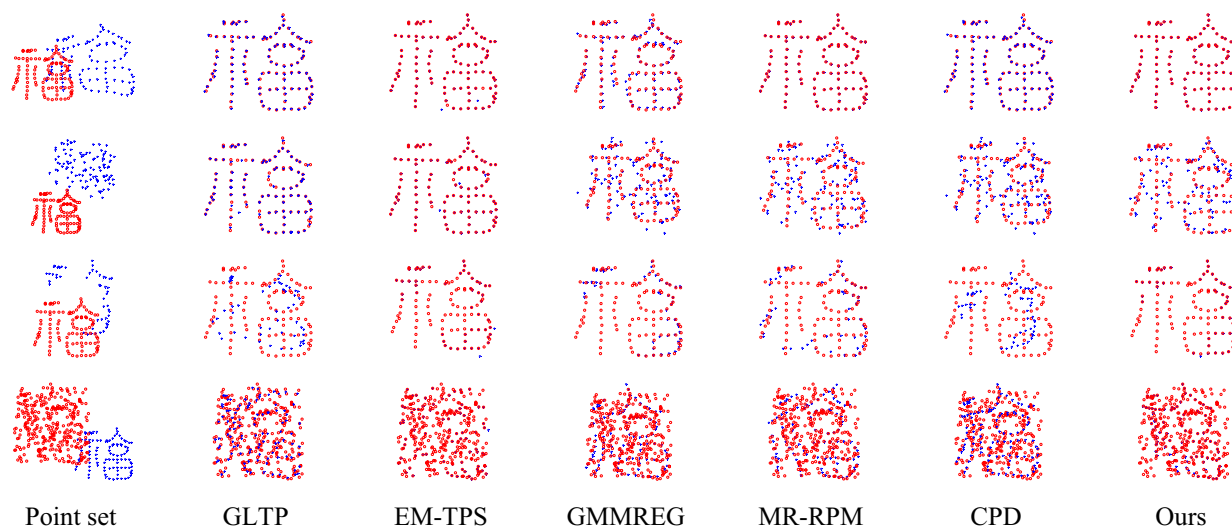

| Point set | GLTP | EM-TPS | GMMREG | MR-RPM | CPD | Ours |

Fig. 3 Point set registration results for *Chinese character* samples with the largest degradations, from the top row to the bottom row: deformation (0.08), noise (0.05), occlusion (0.5), and outliers (2). The leftmost column shows the model point set (blue pluses) and the scene point set (red circles). The goal is to align the model point set with the scene point set. In the second column from the left to the rightmost column, the registration of GLTP, EM-TPS, GMMREG, MR-RPM, CPD, and our method, respectively, is presented
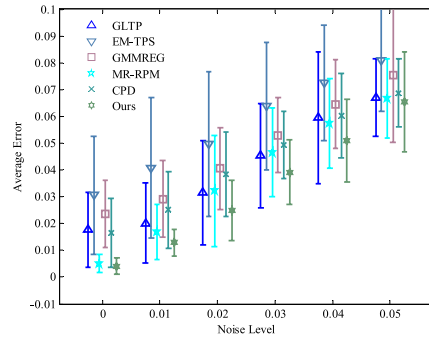
Fig. 4 Comparison of the performance of the registration methods for point sets based on the *Chinese character* shape model with different levels of degradation. **a**, **e**: Deformation test. **b**, **f**: Noise test. **c**, **g**: Occlusion test. **d**, **h**: Outlier test
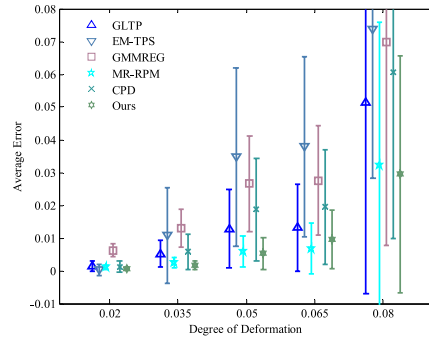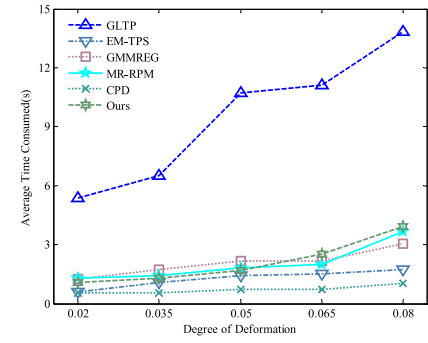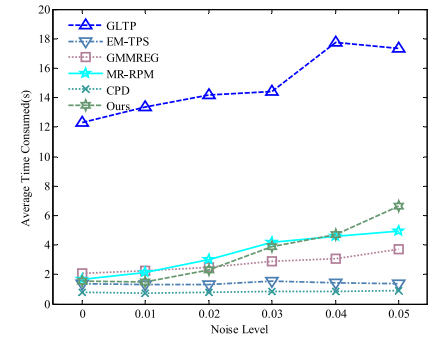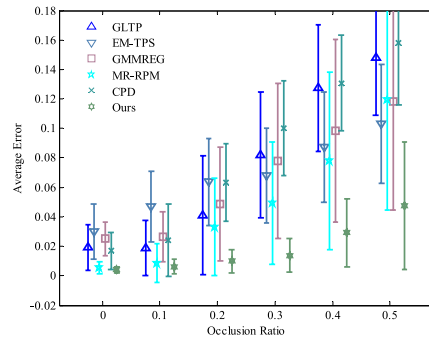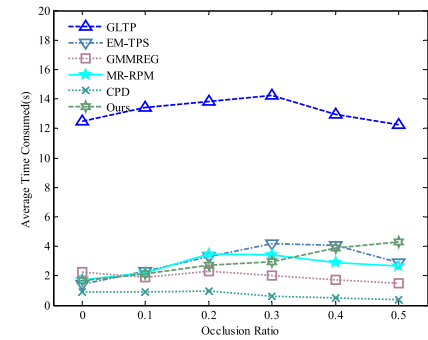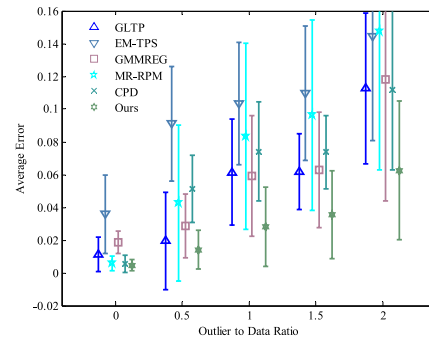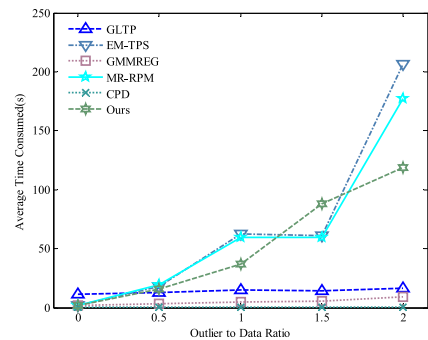
occlusion and outliers and are easily confused at a relatively high degradation level. This is the result of their failure to establish reliable correspondences. In contrast, our proposed method obtain the best registration results because of the correspondences are established subject to the ordering information of the contour points using local descriptors. Figure 4e–h show the average time consumptions for different degradation types and degradation levels. Compared to the *fish* shape, the contour points of the *Chinese character* model are more spread out around the shape, and the shape structure is more complex. Consequently, more time is consumed on the *Chinese character* dataset than on the *fish* dataset. We see that all methods achieve similar time performance on both the *fish* and *Chinese character* datasets, although they need less time to accomplish the registration for each degradation level and each type of degradation in the former case.

In summary, the overall trend of error bars on the *Chinese character* dataset is similar to that on the *fish* dataset. Our proposed method obtains better assignments in almost all cases under all four degradation scenarios. Both the *fish* and *Chinese character* results show that our proposed method is effective for the nonrigid registration of 2D point sets.

4.2 Experimental and analysis on real-world image data

In this subsection, to test and compare the performance of our proposed method on real-world images, we conducted experiments on the Tarragona palmprint [58] and CASIA fingerprint image databases [59]. We collect two groups of image sequences with nonrigid transformation form each database. Here, each sequence contains a series of five homologous images. We randomly choose an image of each sequence is used as the model image, and the others are treated as scene images; thus, we can obtain 4 image pairs from each sequence.

In this experiment, we first constructed the model and scene point sets (150 point pairs) between each pairs of images based on the open source VLFeat toolbox [60], and then the initial correspondences were detected by means of nearest neighbor strategy between the model and scene point sets. Note that due to the limitations of this strategy, not all estimated initial correspondences will be correct. Finally, we aligned the model point set with the scene point set to partition the initial correspondences into inliers and outliers and capture the visual correspondences between the model and scene point sets.

We select several commonly used metrics, namely the accuracy, precision, and recall to evaluate the performance.

The accuracy is the proportion of true results among the total number of point pairs. The precision is the proportion of true inliers among the total number of identified inliers. The recall is the proportion of true inliers among the total number of initial correspondences. Specifically, the accuracy, precision, and recall, which are used to evaluate the feature registration results, are defined as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{21}$$

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \tag{22}$$

$$\text{Recall} = \text{TP}/(\text{TP} + \text{FN}) \tag{23}$$

where the terms TP, FP, TN, and FN denote true positives, false positives, true negatives, and false negatives, respectively.

Under the assumption that the initial correspondences are the ground truth, Fig. 5 and Fig. 6 show the qualitative results of our proposed approach based on the palmprint and fingerprint image databases. Clearly, almost all of the inliers (TP) and outliers (FN) among the initial correspondences are correctly distinguished, and the inliers (TP + FP) between the model and scene point sets can also be established using our approach. According to the experimental results, the performance of our method is better than that of the nearest neighbor strategy. Tables 1, 2, 3, and 4 show the accuracy, precision, and recall for every image sequence, and we see that our method achieves higher accuracy and precision values in most cases, indicating its strong ability to identify inliers. Meanwhile, we note that our method achieves a low recall value on most test image pairs, which means that many outliers are falsely captured among the initial correspondences by the nearest neighbor strategy, and our method can effectively identify outliers in the initial set for each image pair.

Next, we present a quantitative comparison, including the total numbers of identified inliers (TP + FN) and the corresponding average errors between the model and scene point sets. The results are illustrated in Fig. 7. EM-TPS has a better average error than the other methods, followed by our method and MR-RPM; these findings illustrate that these methods achieve more robust results in identifying inliers. However, EM-TPS ignores the importance of preserving the intrinsic geometrical structure information among the point sets and balancing the intrinsic and global constraints during the transformation, which forces disruption of the local structure of the point sets during registration. Thus, the registration results do not generate satisfactory alignments
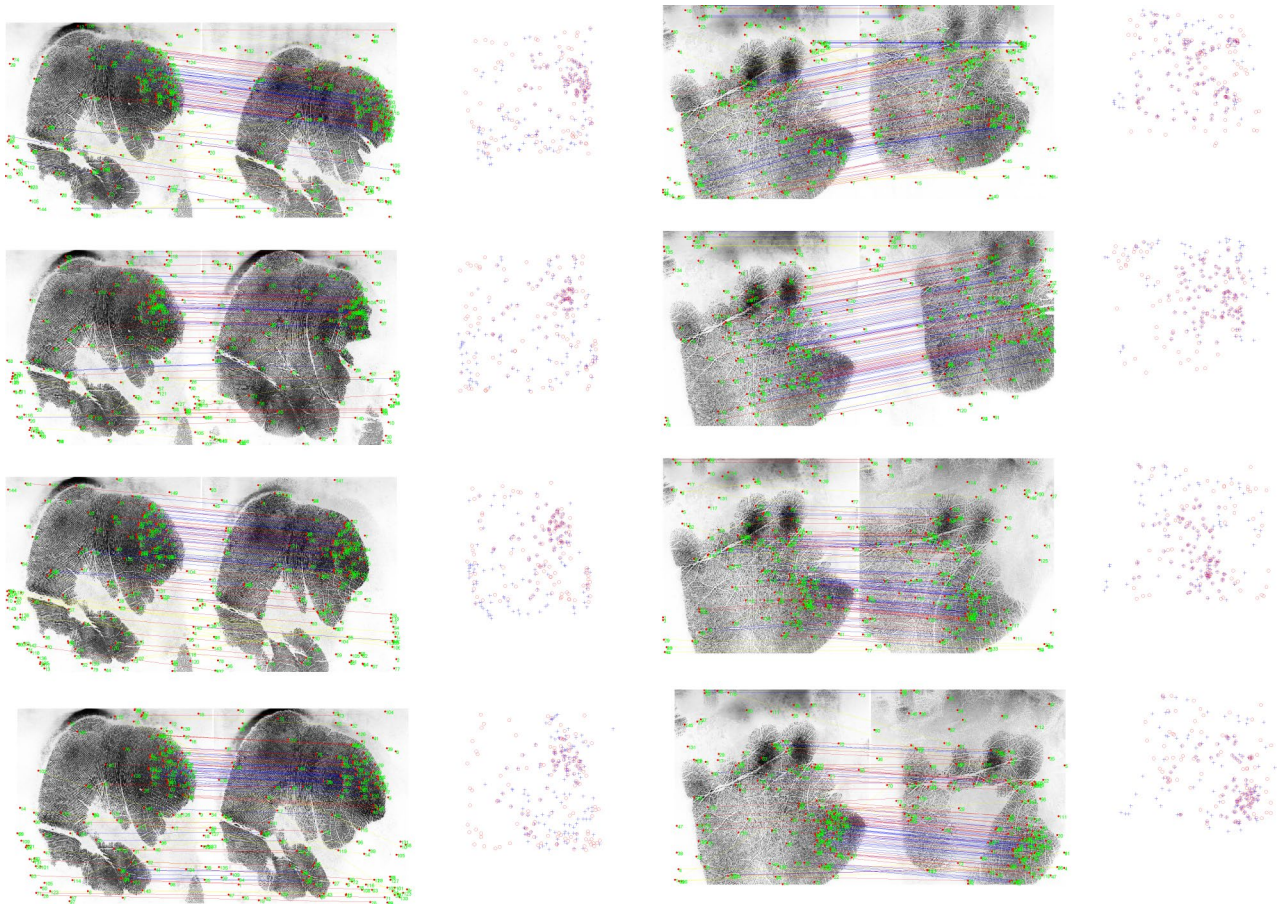
Fig. 5 Qualitative results on right palmprint image pairs (first two columns) and left palmprint image pairs (last two columns) obtained using our method. From the top row to the bottom row, the registration performance is shown for image pairs '1 and 2' to '1 and 5'. The feature point alignment is shown within each set of experiments (the model point set: blue pluses, the scene point set: red circles). The lines indicate the matching results, where the blue (TP) and yellow lines (FN) denote the preserved inliers and removed outliers among the initial correspondences, respectively, and the blue (TP) and red lines (FP) codenote the identified inliers between the model and scene point sets

because some outliers may be falsely identified as inliers, as shown in Fig. 8. Our method and MR-RPM identify similar numbers of inliers. However, considering both accuracy and precision, our method shows slightly better performance.

In addition, the corresponding correct match ratios for every dataset are shown in Figs. 9 and 10. The correct match ratio is the proportion of identified inliers among the total number of point pairs (150 point pairs). As depicted in these two figures, the experimental results of EM-TPS are stable or vary only slightly as the distance level increases. Intuitively, it appears to perform the best due to assigning point matches that capture only the global distributions without considering the intrinsic geometrical structure information, which disturbs the local neighborhood structures of the input data and forces some point pairs between the model and scene point sets achieve accurate alignments. However, the feature correspondences in this method are usually established based on only global spatial relationship, and hence some unknown false matched will be introduced which degrade the homing performance; therefore, this method cannot meet the needs of a real engineering environment.

In some cases, the correct match rates of GLTP indicate that the desired results can be obtained in a distance
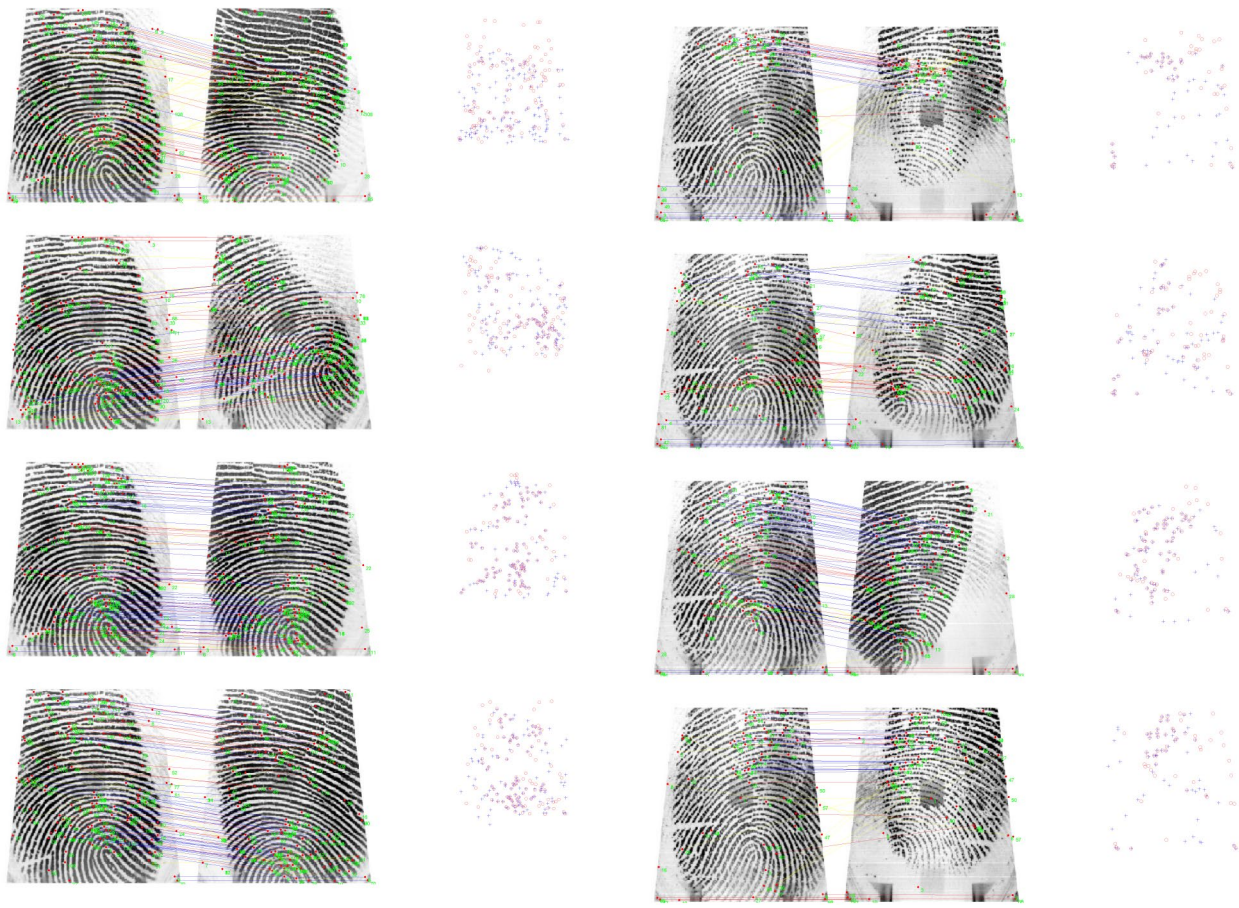
Fig. 6 Qualitative results on right fingerprint image pairs (first two columns) and left fingerprint image pairs (last two columns) obtained using our method. From the top row to the bottom row, the registration performance is shown for image pairs '1 and 2' to '1 and 5'. The feature point alignment is shown within each set of experiments (the model point set: blue pluses, the scene point set: red circles). The lines indicate the matching results, where the blue (TP) and yellow lines (FN) denote the preserved inliers and removed outliers among the initial correspondences, respectively, and the blue (TP) and red lines (FP) codenote the identified inliers between the model and scene point sets

Table 1 Comparison of registration results on right palmprint image datasets in terms of accuracy, precision, and recall (Em: evaluation metric)

Dataset	Em	GLTP	EM-TPS	GMMREG	MR-RPM	CPD	Ours
1 and 2	Accuracy	0.4733	0.6133	0.4667	0.6133	0.4333	**0.6467**
	Precision	0.4312	0.5300	0.4412	0.5385	0.4295	**0.5733**
	Recall	0.7344	0.8281	0.9375	0.6563	1.0000	0.6719
1 and 3	Accuracy	0.3533	0.6200	0.3533	0.6733	0.3000	**0.6800**
	Precision	0.2589	0.4125	0.3043	0.4559	0.2877	**0.4603**
	Recall	0.6744	0.7674	0.9767	0.7209	0.9767	0.6744
1 and 4	Accuracy	0.5267	0.5733	0.4200	**0.6800**	0.3600	0.6600
	Precision	0.4022	0.4396	0.3786	**0.5397**	0.3557	0.5152
	Recall	0.6981	0.7547	1.0000	0.6415	1.0000	0.6415
1 and 5	Accuracy	0.5267	0.5933	0.4000	0.7133	0.3600	**0.7333**
	Precision	0.4100	0.4556	0.3630	0.5735	0.3537	**0.6140**
	Recall	0.7736	0.7736	0.9245	0.7358	0.9811	0.6604

Table 2 Comparison of registration results on left palmprint image datasets in terms of accuracy, precision, and recall (Em: evaluation metric)

Dataset	Em	GLTP	EM-TPS	GMMREG	MR-RPM	CPD	Ours
1 and 2	Accuracy	0.4933	0.6267	0.6067	0.6267	0.5333	**0.6733**
	Precision	0.5047	0.5981	0.5682	0.6180	0.5238	**0.6489**
	Recall	0.7013	0.8312	0.9740	0.7143	1.0000	0.7922
1 and 3	Accuracy	0.5667	0.6467	0.5667	0.6733	0.4733	**0.6867**
	Precision	0.5263	0.5818	0.5231	0.6100	0.4733	**0.6224**
	Recall	0.8451	0.9014	0.9577	0.8592	1.0000	0.8592
1 and 4	Accuracy	0.5400	0.5733	0.5067	0.6067	0.4533	**0.6133**
	Precision	0.4862	0.5102	0.4697	0.5376	0.4459	**0.5444**
	Recall	0.8030	0.7576	0.9394	0.7576	1.0000	0.7424
1 and 5	Accuracy	0.4600	0.5733	0.4267	**0.6200**	0.4200	0.6133
	Precision	0.4018	0.4747	0.3969	**0.5114**	0.4041	0.5059
	Recall	0.7627	0.7966	0.8814	0.7627	1.0000	0.7288

Table 3 Comparison of registration results on right fingerprint image datasets in terms of accuracy, precision, and recall (Em: evaluation metric)

Dataset	Em	GLTP	EM-TPS	GMMREG	MR-RPM	CPD	Ours
1 and 2	Accuracy	0.4000	0.5400	0.4067	0.6067	0.3200	**0.6467**
	Precision	0.3190	0.3636	0.3411	0.3721	0.3200	**0.4419**
	Recall	0.7708	0.5833	0.9167	0.3333	1.0000	0.3958
1 and 3	Accuracy	0.4533	0.5067	0.3867	0.6000	0.3800	**0.6200**
	Precision	0.3832	0.4190	0.3704	0.4815	0.3784	**0.5000**
	Recall	0.7193	0.7719	0.8772	0.6842	0.9825	0.6491
1 and 4	Accuracy	0.5467	0.6467	0.6133	0.7267	0.6133	**0.7333**
	Precision	0.5946	0.6343	0.6069	0.7308	0.6131	**0.7379**
	Recall	0.7416	0.9551	0.9888	0.8539	0.9438	0.8539
1 and 5	Accuracy	0.63.33	0.6533	0.5800	0.7133	0.5400	**0.7200**
	Precision	0.5794	0.5943	0.5344	0.6629	0.5106	**0.6667**
	Recall	0.8611	0.8750	0.9722	0.8194	1.0000	0.8333

Table 4 Comparison of registration results on left fingerprint image datasets in terms of accuracy, precision, and recall (Em: evaluation metric)

| Dataset	Em	GLTP	EM-TPS	GMMREG	MR-RPM	CPD	Ours
1 and 2	Accuracy	0.5538	0.6154	0.6154	**0.6923**	0.4923	0.6769
	Precision	0.5349	0.5854	0.5686	**0.6667**	0.4923	0.6486
	Recall	0.7188	0.7500	0.9063	0.7500	1.0000	0.7500
1 and 3	Accuracy	0.5244	0.6341	0.4634	0.6463	0.4634	**0.6585**
	Precision	0.4808	0.5614	0.4478	0.6111	0.4557	**0.6154**
	Recall	0.6757	0.8649	0.8108	0.5946	0.9730	0.6486
1 and 4	Accuracy	0.6400	0.7500	0.7700	0.7600	0.7500	**0.7800**
	Precision	0.7183	0.7500	0.7558	0.7867	0.7561	**0.7922**
	Recall	0.7612	0.9403	0.9701	0.8806	0.9254	0.9104
1 and 5	Accuracy	0.4638	0.5362	0.5797	0.6232	0.4928	**0.6812**
	Precision	0.4694	0.5200	0.5490	0.6111	0.4928	**0.6667**
	Recall	0.6765	0.7647	0.8235	0.6471	1.0000	0.7059
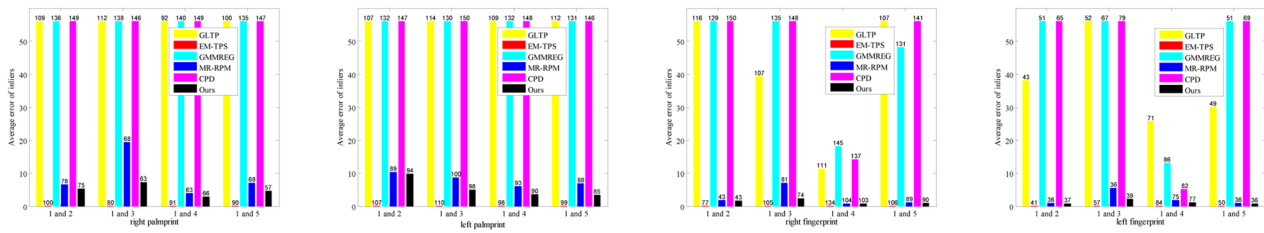
Fig. 7 Comparison of results on real image datasets. The total numbers of identified inliers are shown on the histogram end caps, and the average errors are plotted
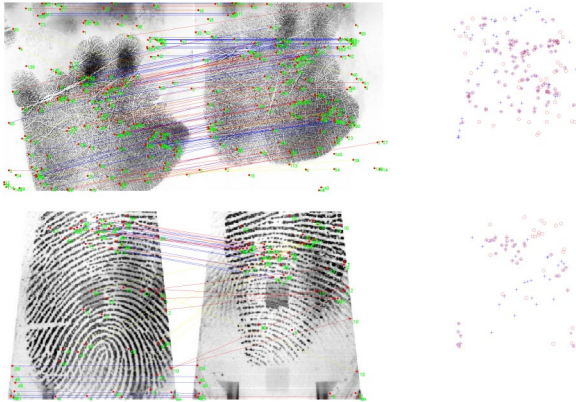

Fig. 8 Qualitative results on left palmprint image pairs '1 and 2' (top row) and left fingerprint image pairs '1 and 2' (bottom row) obtained using EM-TPS

range of 0–4. However, there are fewer captured correspondences as the distance between the corresponding features increases. The GMMREG and CPD methods fail due to the input data are contaminated by outliers. In contrast, MR-RPM and our method achieve better objective performance among the compared approaches. Overall, they can establish more reliable correspondences than the other methods.

5 Conclusion

In this paper, we propose a novel approach for nonrigid point set registration, in which the registration problem is formulated as a probabilistic model that is solved using an iterative EM method. A key characteristic of our approach is that it incorporates a TPS-based constraint and a spectral graph-based constraint into a unified GMM-based framework to preserve the global and intrinsic topological characteristics of the point sets after transformation. In addition, the spatial transformation related to the model point sets is parameterized using the TPS approach, allowing the transformation to be decomposed into affine and nonaffine components. This TPS-based minimization framework has the advantages of free parameters and closed-form solutions in the cases of different degradations. We experimentally tested our approach on contour image and real image datasets and compared it to GLTP, EM-TPS, GMMREG, MR-RPM, and CPD, and the results demonstrate the superiority of our proposed approach in most scenarios.
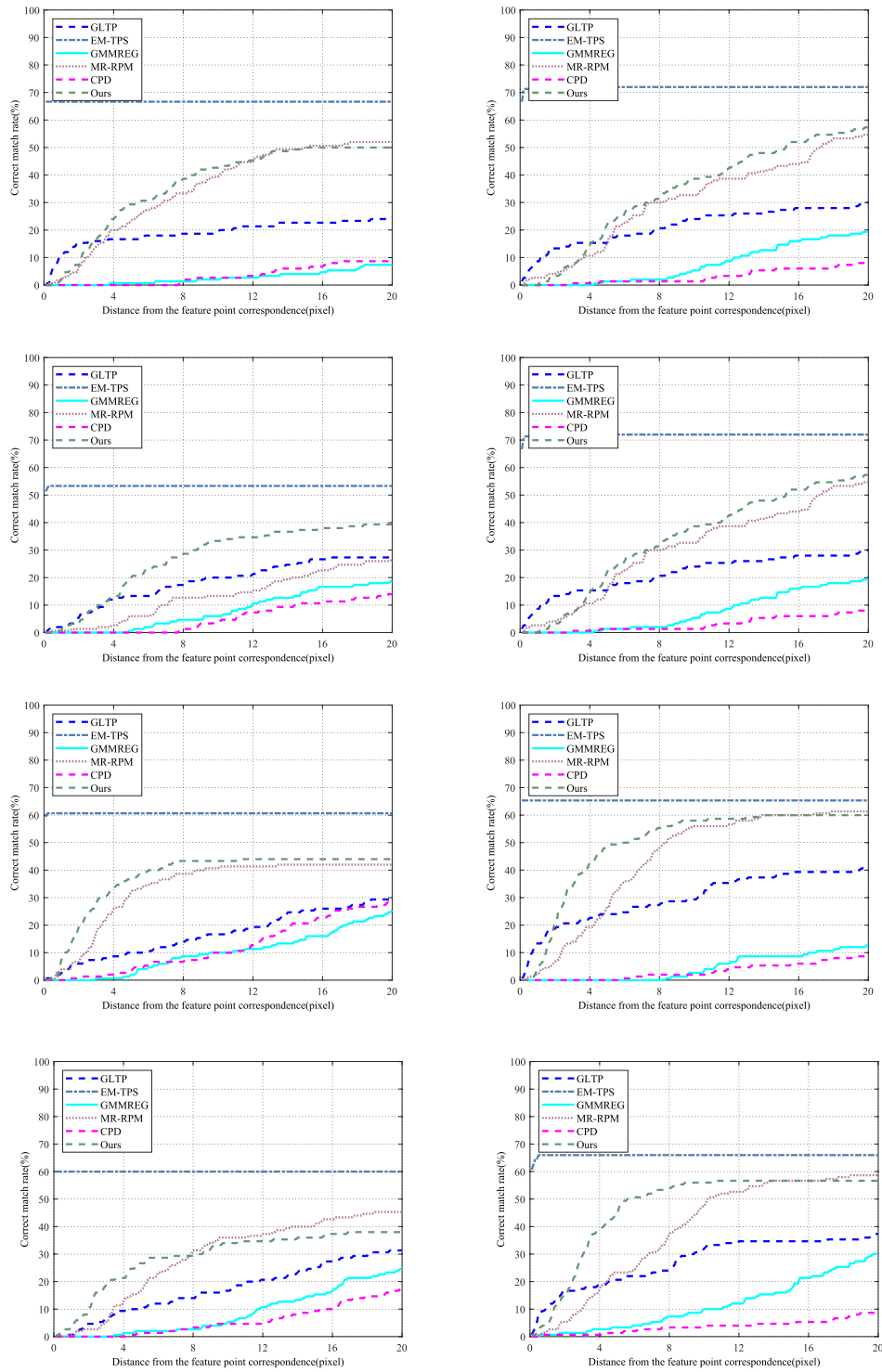
Fig. 9 Performance of all tested methods on right palmprint (right column) and left palmprint (left column) datasets in terms of the correct match ratio
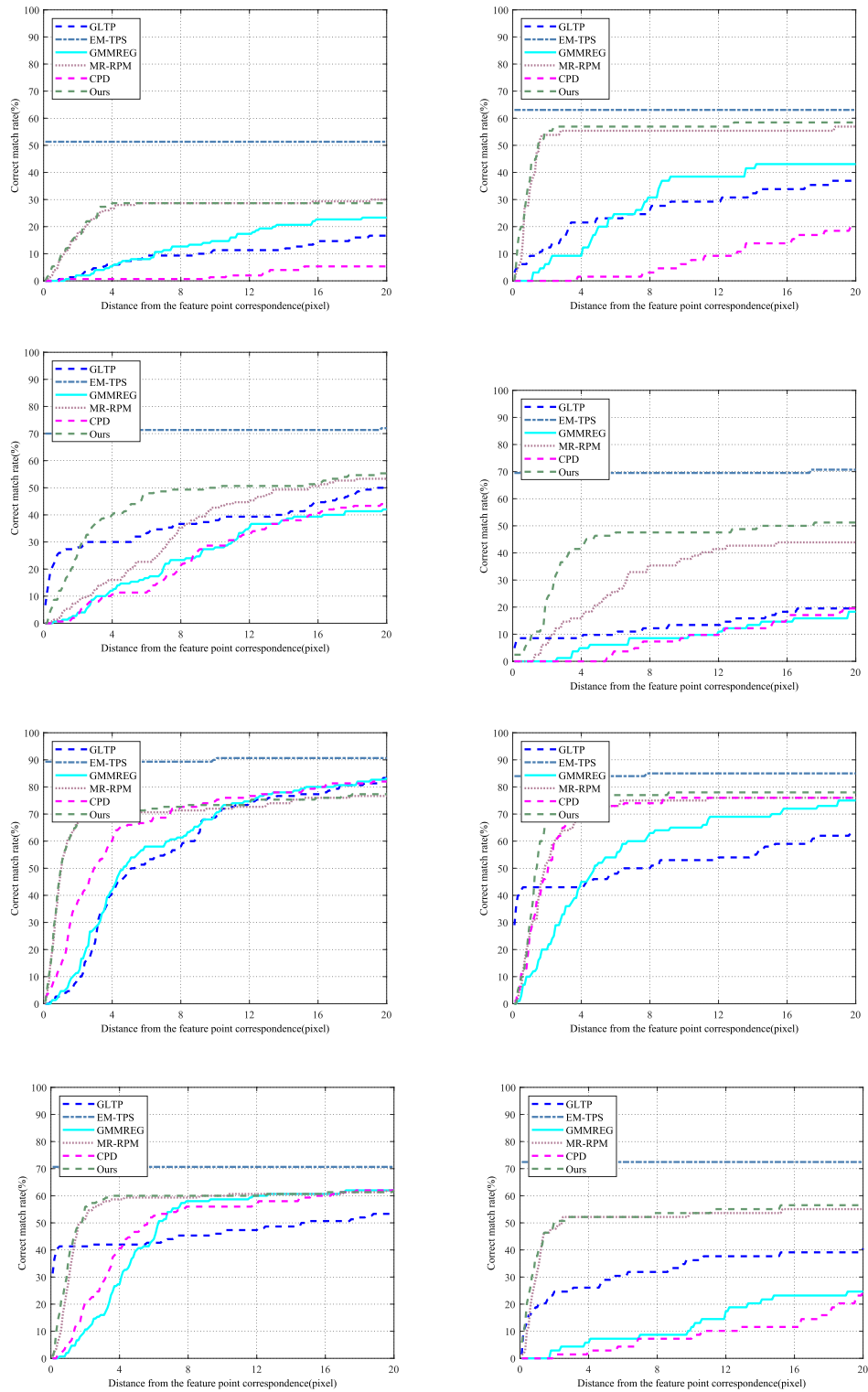
Fig. 10 Performance of all tested methods on right fingerprint (right column) and left fingerprint (left column) datasets in terms of the correct match ratio

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

References

1. De Vos, B.D., Berendsen, F.F., Viergever, M.A., Sokooti, H., Staring, M., Isgum, I.: A deep learning framework for unsupervised affine and deformable image registration. Med. Image Anal. **52**, 128–143 (2019)

2. Ma, J., Zhao, J., Jiang, J., Zhou, H., Guo, X.: Locality preserving matching. Int. J. Comput. Vis. **127**(5), 512–531 (2019)

3. Hu, L., Xiao, J., Wang, Y.: An automatic 3D registration method for rock mass point clouds based on plane detection and polygon matching. Vis. Comput. **36**, 669–691 (2020)

4. Krishnakumar, K., Gandhi, S.I.: Video stitching based on multi-view spatiotemporal feature points and grid-based matching. Vis. Comput. **36**, 1837–1846 (2020)

5. Kan, P., Kaufmann, H.: Deeplight: light source estimation for augmented reality using deep learning. Vis. Comput. **35**(6–8), 873–883 (2019)

6. Choi, J., Son, M.G., Lee, Y.Y., Lee, K.H., Park, J.P., Yeo, C.H., Park, J.S., Choi, S.G., Kim, W.D., Kang, T.W., Ko, K.H.: Position-based augmented reality platform for aiding construction and inspection of offshore plants. Vis. Comput. (2020). https://doi.org/10.1007/s00371-020-01902-9

7. De Lima, R., Cabreraponce, A.A., Martinezcarranza, J.: Parallel hashing-based matching for real-time aerial image mosaicing. J. Real-time Image Process. (2020). https://doi.org/10.1007/s11554-020-00959-y

8. Lati, A., Belhocine, M., Achour, N.: Robust aerial image mosaicing algorithm based on fuzzy outliers rejection. Evolv. Syst. (2019). https://doi.org/10.1007/s12530-019-09279-4

9. Choy, C., Lee, J., Ranftl, R., Park, J., Koltun, V.: High-dimensional convolutional networks for geometric pattern recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11227–11236 (2020)

10. Hammouda, G., Sellami, D., Hammouda, A.: Pattern recognition based on compound complex shape-invariant Radon transform. Vis. Comput. **36**(2), 279–290 (2020)

11. Ullah, K., Mahmood, T., Ali, Z., Jan, N.: On some distance measures of complex Pythagorean fuzzy sets and their applications in pattern recognition. Complex Intell. Syst. **6**(1), 15–27 (2020)

12. Yang, C., Feng, H., Xu, Z., Li, Q., Chen, Y.: Correction of overexposure utilizing haze removal model and image fusion technique. Vis. Comput. **35**(5), 695–705 (2019)

13. Besl, P.J., Mckay, H.D.: A method for registration of 3-D shape. IEEE Trans. Pattern Anal. Mach. Intell. **14**(2), 239–256 (1992)

14. Brown, L.G.: A survey of image registration techniques. ACM Comput. Surv. (CSUR) **24**(4), 325–376 (1992)

15. Makadia, A., Patterson, A., Daniilidis, K.: Fully automatic registration of 3D point clouds. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1297–1304 (2006)

16. Zhang, S., Wang, H., Huang, W.: Palmprint identification combining hierarchical multi-scale complete LBP and weighted SRC. Soft. Comput. **24**(6), 4041–4053 (2020)

17. An, F., Liu, Z.: Facial expression recognition algorithm based on parameter adaptive initialization of CNN and LSTM. Vis. Comput. **36**(3), 483–498 (2020)

18. Krejsa, J., Vechet, S.: Evaluation of visual markers detection used for autonomous mobile robot docking navigation. In: International Conference Mechatronics, pp. 229–236 (2019)

19. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)

20. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002)

21. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. b-Methodol. **39**(1), 1–22 (1977)

22. Wahba, G.: Spline models for observational data. Siam (1990)

23. Rusu, R. B., Blodow, N., Beetz, M.: Fast point feature histograms (FPFH) for 3D registration. In: IEEE International Conference on Robotics and Automation, pp. 3212–3217 (2009)

24. Hauagge, D.C., Snavely, N.: Image matching using local symmetry features. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 206–213 (2012)

25. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. **27**(10), 1615–1630 (2005)

26. Rusinkiewicz, S., Levoy, M.: Efficient variants of the ICP algorithm. In: Proceedings Third International Conference on 3-D Digital Imaging and Modeling, pp. 145–152 (2001)

27. Amberg, B., Romdhani, S., Vetter, T.: Optimal step nonrigid ICP algorithms for surface registration. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2007)

28. Zhu, J., Du, S., Yuan, Z., Liu, Y., Ma, L.: Robust affine iterative closest point algorithm with bidirectional distance. IET Comput. Vis. **6**(3), 252–261 (2012)

29. Pomerleau, F., Colas, F., Siegwart, R., Magnenat, S.: Comparing ICP variants on real-world data sets. Auton. Robots. **34**(3), 133–148 (2013)

30. Chui, H., Rangarajan, A.: A new point matching algorithm for non-rigid registration. Comput. Vis. Image Underst. **89**(2), 114–141 (2003)

31. Tsin, Y., Kanade, T.: A correlation-based approach to robust point set registration. In: European Conference on Computer Vision, pp. 558–569 (2004)

32. Jian, B., Vemuri, B.C.: A robust algorithm for point set registration using mixture of Gaussians. In: International Conference on Computer Vision, pp. 1246–1251 (2005)

33. Yuille, A.L., Grzywacz, N.M.: The motion coherence theory. In: International Conference on Computer Vision, pp. 344–353 (1988)

34. Yuille, A.L., Grzywacz, N.M.: A mathematical analysis of the motion coherence theory. Int. J. Comput. Vis. **3**(2), 155–175 (1989)

35. Myronenko, A., Song, X., Carreiraperpinan, M.A.: Non-rigid point set registration: coherent point drift. In: Advances in Neural Information Processing Systems, pp. 1009–1016 (2007)

36. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. **32**(12), 2262–2275 (2010)

37. Yang, Y., Ong, S.H., Foong, K.W.C.: A robust global and local mixture distance based non-rigid point set registration. Pattern Recognit. **48**(1), 156–173 (2015)

38. Ma, J., Zhao, J., Yuille, A.L.: Non-rigid point set registration by preserving global and local structures. IEEE Trans. Image Process. **25**(1), 53–64 (2016)

39. Yang, C., Liu, Y., Jiang, X., Zhang, Z., Wei, L., Lai, T., Chen, R.: Non-rigid point set registration via adaptive weighted objective function. IEEE Access. **6**, 75947–75960 (2018)

40. Wang, G., Zhou, Q., Chen, Y.: Robust non-rigid point set registration using spatially constrained Gaussian fields. IEEE Trans. Image Process. **26**(4), 1759–1769 (2017)

41. Ling, H., Jacobs, D.W.: Shape classification using the inner-distance. IEEE Trans. Pattern Anal. Mach. Intell. **29**(2), 286–299 (2007)

42. Ma, J., Zhao, J., Jiang, J., Zhou, H.: Non-rigid point set registration with robust transformation estimation under manifold regularization. In: Thirty-First AAAI Conference on Artificial Intelligence, pp. 4218–4224 (2017)

43. Ma, J., Wu, J., Zhao, J., Jiang, J., Zhou, H., Sheng, Q.Z.: Non-rigid point set registration with robust transformation learning under manifold regularization. IEEE Trans. Neural Netw. **30**(12), 3584–3597 (2019)

44. Wang, J., Chen, J., Xu, H., Zhang, S., Mei, X., Huang, J., Ma, J.: Gaussian field estimator with manifold regularization for retinal image registration. Sig. Process. **157**, 225–235 (2019)

45. Ge, S., Fan, G.: Non-rigid articulated point set registration with local structure preservation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 126–133 (2015)

46. Ge, S., Fan, G.: Topology-aware non-rigid point set registration via global-local topology preservation. Mach. Vis. Appl. **30**(4), 717–735 (2019)

47. Ma, J., Jiang, J., Zhou, H., Zhao, J., Guo, X.: Guided locality preserving feature matching for remote sensing image registration. IEEE Trans. Geosci. Remote Sens. **56**(8), 4435–4447 (2018)

48. Ma, J., Jiang, X., Jiang, J., Zhao, J., Guo, X.: LMR: learning a two-class classifier for mismatch removal. IEEE Trans. Image Process. **28**(8), 4045–4059 (2019)

49. Sedaghat, A., Mohammadi, N.: High-resolution image registration based on improved SURF detector and localized GTM. Int. J. Remote Sens. **40**(7), 2576–2601 (2019)

50. Ma, J., Jiang, X., Jiang, J., Zhao, J., Guo, X.: Feature-guided Gaussian mixture model for image matching. Pattern Recognit. **92**, 231–245 (2019)

51. Jiang, X., Ma, J., Jiang, J., Guo, X.: Robust feature matching using spatial clustering with heavy outliers. IEEE Trans. Image Process. **29**, 736–746 (2019)

52. Xiao, D., Zahra, D., Bourgeat, P., Berghofer, P., Tamayo, O.A., Wimberley, C., Gregoire, M.C., Salvado, O.: An improved 3D shape context based non-rigid registration method and its application to small animal skeletons registration. Comput. Med. Imaging Graph. **34**(4), 321–332 (2010)

53. Deng, W., Zou, H., Guo, F., Lei, L., Zhou, S., Luo, T.: A robust non-rigid point set registration method based on inhomogeneous Gaussian mixture models. Vis. Comput. **34**(10), 1399–1414 (2018)

54. Schaefer, S., McPhail, T., Warren, J.: Image deformation using moving least squares. In: ACM SIGGRAPH 2006 Papers, pp. 533–540 (2006)

55. Belkin, M., Niyogi, P.: Laplacian eigenmaps and spectral techniques for embedding and clustering. In: Advances in Neural Information Processing Systems, pp. 585–591 (2002)

56. Belkin, M., Niyogi, P., Sindhwani, V.: Manifold regularization: a geometric framework for learning from labeled and unlabeled examples. J. Mach. Learn. Res. **7**(11), 2399–2434 (2006)

57. Yang, G., Li, R., Liu, Y., Wang, J.: A unified framework for non-rigid point set registration via coregularized least squares. IEEE Access **8**, 130263–130280 (2020)

58. Moreno-Garcia, C.F., Serratosa, F.: Correspondence consensus of two sets of correspondences through optimisation functions. Pattern Anal. Appl. **20**(1), 201–213 (2017)

59. Galbally, J., Alonso-Fernandez, F., Fierrez, J., Ortega-Garcia, J.: A high performance fingerprint liveness detection method based on quality related features. Future Generat. Comput. Syst. **28**(1), 311–321 (2012)

60. Vedaldi, A., Fulkerson, B.: VLFeat: An open and portable library of computer vision algorithms. In: Proceedings of the 18th ACM International Conference on Multimedia, pp. 1469–1472 (2010)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Guiqiang Yang received the B.E. degree in mechanical design manufacture and automation from Wuhan Textile University, Wuhan, China, in 2010 and the M.E. degree in design and construction of naval architecture and ocean structure from JiMei University, Xiamen, China, in 2013. He is currently working toward a Ph.D. degree in the School of Naval Architecture at Dalian University of Technology (DLUT), China. His main research directions are intelligent manufacturing, machine learning and data analysis, and robust model fitting.

Rui Li received the B.E. and Ph.D. degrees in design and construction of naval architecture and ocean structure from Dalian University of Technology (DLUT), Dalian, China, in 2004 and 2010, respectively. Since 2015, he has been an Associate Professor and the M.E. Supervisor with the College of Shipbuilding Engineering, Dalian University of Technology. He has published more than 30 papers. His research interests include intelligent manufacturing, machine learning and data analysis, and data registration.

Yujun Liu received the Ph.D. degree from Dalian University of Technology (DLUT), Dalian, China, in 1997. Since 1999, he has been a Professor in College of Shipbuilding Engineering, Dalian University of Technology. He enjoyed the Special Allowance of the State Council in 2001. His current research interests include digital shipbuilding, data analysis and visualization, and data registration.

Ji Wang received the B.E. and Ph.D. degrees in design and construction of naval architecture and ocean structure from Dalian University of Technology (DLUT), Dalian, China, in 2000 and 2007, respectively. He is currently a Professor with the College of Shipbuilding Engineering, Dalian University of Technology. His research interests include digital shipbuilding, data analysis and processing, and intelligent manufacturing.