

Context-Aware Gaussian Fields for Non-rigid Point Set Registration

Gang Wang, Zhicheng Wang, Yufei Chen, Qiangqiang Zhou, Weidong Zhao
 CAD Research Center, College of Electronics and Information Engineering
 Tongji University, Shanghai 201804, China

gwang.cv@gmail.com, {zhichengwang, yufeichen, wd}@tongji.edu.cn

Abstract

Point set registration (PSR) is a fundamental problem in computer vision and pattern recognition, and it has been successfully applied to many applications. Although widely used, existing PSR methods cannot align point sets robustly under degradations, such as deformation, noise, occlusion, outlier, rotation, and multi-view changes. This paper proposes context-aware Gaussian fields (CA-LapGF) for non-rigid PSR subject to global rigid and local non-rigid geometric constraints, where a laplacian regularized term is added to preserve the intrinsic geometry of the transformed set. CA-LapGF uses a robust objective function and the quasi-Newton algorithm to estimate the likely correspondences, and the non-rigid transformation parameters between two point sets iteratively. The CA-LapGF can estimate non-rigid transformations, which are mapped to reproducing kernel Hilbert spaces, accurately and robustly in the presence of degradations. Experimental results on synthetic and real images reveal that how CA-LapGF outperforms state-of-the-art algorithms for non-rigid PSR.

1. Introduction

Point set registration (PSR) has been widely applied in computer vision and pattern recognition to solve many problems such as robot navigation, motion tracking, biomedical image registration [25], and face recognition. The main purpose of the registration problem is to estimate likely correspondences between two point sets and recover the underlying transformation which can align the corresponding point pairs perfectly.

Generally speaking, point set registration can be categorized into rigid and non-rigid depending on the transformation pattern. The former one is relatively easy to estimate, while the non-rigid transformation is hard to estimate due to the underlying transformation model is usually unknown and difficult to approximate. Moreover, as a key component in point set registration, the non-rigid transformation exists in numerous applications, including hand-written character

recognition, and facial-expression recognition.

However, point set registration becomes increasingly difficult, because of some challenges: (1) the sensitive registration accuracy in the presence of large degree of degradations such as deformation, noise, occlusion, outlier, rotation, and multi-view changes, which make the distribution of point set more complex, where noisy data means the feature points cannot be matched precisely, and the data with occlusion and outliers mean some points cannot find their underlying correspondences in the corresponding point set; (2) the numerical optimization often falls into local minima; (3) the high computational complexity when handling a large number of points.

In face of these challenges, numerous registration algorithms have been proposed recently. The Iterative Closest Point (ICP) algorithm [3] is with simplicity and low computational complexity, which uses the nearest-neighbor distance criterion to assign binary correspondences, and the least squares to estimate the rigid transformation iteratively. However, ICP requires an initial position, such as an adequately close distance between two point sets.

For non-rigid transformation, Chui and Rangarajan [6] introduced a soft assignment technique and the deterministic annealing to construct a general framework to estimate the fuzzy correspondences and recover the non-rigid transformation parameterized by Thin Plate Spline (TPS) iteratively. Then a robust point set registration algorithm (TPS-RPM) has been presented. Although it is more robust than ICP when confronting some degree of degradations, it is with high computational complexity. Zheng et al. [30] proposed a robust point matching by preserving local neighborhood structures (RPM-PLNS) for non-rigid shape registration, where graph matching technique is used to preserve local neighborhood structures, but it is still sensitive to outliers. Kernel correlation (KC) [22] considers the correlation between two point set kernel densities, where the underlying transformation parameters can be estimated by maximizing the correlation based on the M-estimator. Based on the theory of the KC, a robust point set registration approach using Gaussian mixture models (GMMReg) [10] has been

presented, it leverages the closed-form expression for the L_2 distance between two Gaussian mixtures which are used to represent the given point sets.

Based on the motion coherence theory (MCT) [28], Myronenko et al. [20] constructed a mixture model, and then proposed an efficient registration algorithm, namely coherence point drift (CPD), where one of the two point sets is modeled as a GMM, the other is the data point set, and the correspondence problem is formulated as a density estimation problem. More precisely, Expectation-Maximization (EM) algorithm is used to solve this mixture model, and the Gaussian radial basis function (GRBF) is used to build the transformation model instead of the TPS for non-rigid transformations. Although, it can handle a large number of points with fast Gaussian transformation (FGT) [9], the CPD needs to estimate the underlying number of Gaussian components, and it is sensitive to occlusion and outliers.

Moreover, Li et al. [11] proposed an asymmetric shape representation and a new high-peak-fat-tail Gaussian mixtures kernel method to align two shapes, and the particle swarm optimization (PSO) is applied to recover the optimal transformation parameters instead of the gradient-based algorithms. Ma et al. [18, 15] introduced a robust estimator in statistics, namely L_2 -minimizing estimate (L_2 E), to estimate the non-rigid transformation, then they proposed a robust point matching algorithm based on L_2 E (RPM- L_2 E), where it needs putative correspondences estimated by shape context descriptor for non-rigid point set registration. A non-rigid point set registration method based on asymmetric Gaussian representation [27, 24] uses a mixture of asymmetric Gaussians to represent point sets, and it updates correspondences and transformations under the framework of TPS-RPM. Another interesting fast point matching method uses a quadratic programming based cluster correspondence projection (QPCCP) [13], but it is sensitive to degradations.

In this paper, we focus on the non-rigid transformation and introduce a robust non-rigid point set registration algorithm: context-aware Gaussian fields (CA-GF). The proposed CA-GF tries to address the registration problem accurately and robustly under the aforementioned degradations, and to overcome the limitations of the existing algorithms. Briefly, the key idea of our method is to find meaningful correspondences using the context information of points by computing their inner distances, and to estimate the underlying non-rigid transformation using robust point matching by Gaussian fields. More specially, inner distance based context-aware strategy is used to represent point sets, and to estimate the likely correspondences, where the inner distance represents the shortest path distance between points within the point set outermost silhouette [14]. Due to the estimated correspondence is designed as an attribute weight of the Gaussian fields, then robust estimation of non-rigid transformation can be obtained. Based on the properties of

reproducing kernel Hilbert spaces (RKHS) with the Representer theorem, non-rigid transformations can be mapped into the RKHS, and the regularization framework is applied to let them become smooth and well defined. Considering the intrinsic geometry of the transformed point sets, the Laplacian regularization term is added into the objective function, and then we call the context-aware Laplacian regularized Gaussian fields (CA-LapGF) method. Under the determined annealing framework, the objective function can be optimized by the quasi-Newton technique. Moreover, low-rank kernel matrix approximation is applied to reduce runtime when facing large number of points. Extensive experiments on some synthesize and real image datasets demonstrate that both CA-GF and CA-LapGF are more robust in the presence of a large degree of degradations (deformation, noise, occlusion, outlier, rotation and multi-view changes). Further more, the proposed robust Gaussian fields algorithm can be well applied to remove mismatches for robust point matching. Comparing with existing Gaussian fields based methods, Gaussian fields framework [4] is used to register three-dimensional rigid surface, and then [16] applied it to register non-rigid visible and infrared face images with adding the Tikhonov regularization theory, while the context-aware Gaussian fields algorithm mainly focuses on the non-rigid transformation, and is applied to register non-rigid point sets and remove mismatches from a putative matching with more accuracy and robustness.

Briefly, the main contributions of our work includes: 1) we use the inner distance based context representation strategy to estimate fuzzy correspondences instead of kernel density estimation and soft-assignment, because of its insensitivity when facing some variety of deformations; 2) iterative updating strategy between correspondences and transformations let us simplify the Gaussian mixture models to the Gaussian fields which can estimate the non-rigid transformations robustly; 3) we apply the robust Gaussian fields algorithm to non-rigid transformation and mismatch removal, and the experiments demonstrate that the algorithm outperforms other registration methods in state-of-the-art.

2. Method

2.1. Problem Formulation

Given two point sets (see notation¹), the model set $\mathbf{X} = \{x_i | x_i \in \mathbb{R}^d, i = 1, \dots, N\}$ and the scene set $\mathbf{Y} = \{y_j | y_j \in \mathbb{R}^d, j = 1, \dots, M\}$, where the scene

¹Bold capital letters denote a matrix \mathbf{X} , x_i denotes the i^{th} row of the matrix \mathbf{X} . x_{ij} denotes the scalar value in the i^{th} row and j^{th} column of the matrix \mathbf{X} . $\mathbf{1}_{m \times n}$ denotes a matrix with all ones, as well as $\mathbf{0}_{m \times n}$ denotes a matrix with all zeros. $\mathbf{I}_{n \times n} \in \mathbb{R}^{n \times n}$ denotes an identity matrix. $\|\cdot\|$ denotes a 2-norm. $\text{tr}(\mathbf{X})$ denotes the trace of the matrix. $\text{diag}(x)$ is a diagonal matrix whose diagonal elements are x . $\mathbf{X} \circ \mathbf{Y}$ is the Hadamard product of matrices, and $\mathbf{X} \otimes \mathbf{Y}$ is the Kronecker product of matrices.

set \mathbf{Y} is fixed as a target set, and the model set \mathbf{X} is moving onto the scene set by a series of transformations $\mathcal{T} = \{\tau_k(\cdot)\}_{k=1}^{N_{iter}}$ where N_{iter} is the number of iteration process, $\tau(\cdot) : \mathbb{R}^d \mapsto \mathbb{R}^d$ denotes a spatial transformation for displacement field and it is parameterized by \mathcal{T} .

We simplify the Gaussian mixture model of Gaussian fields with context information by assuming that the number of Gaussian components can be estimated. Thus the context-aware Gaussian fields (CA-GF) can be formulated for point set registration, and the CA-GF aims to estimate the likely correspondences and the underlying spatial transformation between points such that the sum of distances is minimized:

$$\begin{aligned} & \underset{\mathcal{C}, \mathcal{T}}{\operatorname{argmin}} \mathcal{G}(\mathcal{C}, \mathcal{T}) = \\ & \sum_{i,j} \exp \left(-\frac{c_{ij}\|y_j - x_i - \tau(x_i)\|^2}{\sigma^2} \right) + \phi(\tau), \quad (1) \\ & \text{s.t. } \mathcal{C} \in \Pi, \tau \in \mathcal{H} \end{aligned}$$

where \mathcal{C} denotes the context-aware weight to assign correspondences between points will be discussed in the next subsection, the algorithm proceeds by optimizing between \mathcal{C} and \mathcal{T} in an alternating fashion, Π denotes a permutation matrix in each iteration, \mathcal{H} is a reproducing kernel Hilbert space (RKHS). Note that the non-rigid transformation is mapped to a special feature space \mathcal{H} , since the properties of RKHS with the Representer theorem provide a theoretical basis for the CA-GF algorithm.

2.2. Context-Aware Strategy

In order to estimate the underlying Gaussian components of the Gaussian fields, context-aware strategy is introduced to find the correspondence between points. For registration problem, the assignment matrix \mathcal{C} is either a one-to-one or many-to-one mapping, and is either a soft-assignment with probability or a hard-assignment with $\{0, 1\}$. Here, let \mathcal{C} be a one-to-one hard-assignment:

$$\mathcal{C}(x_i, y_j) = c_{ij} = \begin{cases} 0, & y_j \neq x_i + \tau(x_i), \text{ unmatch} \\ 1, & y_j = x_i + \tau(x_i), \text{ match} \end{cases} \quad (2)$$

where \mathcal{C} can be estimated by the inner distance based context descriptor [14].

Precisely, the shape context (SC) [2] is used to describe the relative spatial distribution of positions around the certain feature points which need to be represented. For instance, the context at point x_i is described by a histogram h_i of the relative coordinates of the remaining $n - 1$ points

$$h_i(k) = \#\{x_j \neq x_i : \delta(x_j, x_i) \in \text{bin}(k)\}, \quad (3)$$

where $\delta(x_j, x_i)$ denotes the inner distance between two points, and the bins are uniform in the log-polar space. As

suggested in [14], instead of the Euclidean distance, the inner distance captures better shape structure, and offers more discriminability for complex shape point sets. Then the match cost between x_i and y_j can be measured by their K -bin normalized context histograms, $h_i(k)$ and $h_j(k)$ respectively, using the Chi-squared test statistic. The assignment algorithm (standard dynamical programming (DP)) has $\mathcal{O}(N^3)$ complexity, however, DP costs $\mathcal{O}(N^2)$ in the method under an ordering constraint on the contour points [14]. DP is more efficient and accurate since it uses the ordering information provided by shape contours. In this paper, dynamical programming is used to match point set \mathbf{X} and \mathbf{Y} in $\mathcal{O}(N^2)$ runtime instead of the Hungarian method in $\mathcal{O}(N^3)$, and then we can get the context-aware weight \mathcal{C} .

2.3. Laplacian Regularized Gaussian Fields

In this subsection, we introduce our Laplacian regularized Gaussian fields (LapGF) algorithm, which preserves the intrinsic geometry of the moving model point set. Briefly, under the manifold regularization framework [1], an additional penalty regularization term is added to penalize the transformation τ along a low dimensional manifold. Thus the Eq. 1 can be rewritten as

$$\begin{aligned} & \underset{\mathcal{C}, \mathcal{T}}{\operatorname{argmin}} \mathcal{G}(\mathcal{C}, \mathcal{T}) = \\ & \sum_{i,j} \exp \left(-\frac{c_{ij}\|y_j - x_i - \tau(x_i)\|^2}{\sigma^2} \right) + \lambda_1 \|\tau\|_{\mathcal{H}}^2 + \lambda_2 \|\tau\|_{\mathcal{M}}^2, \\ & \text{s.t. } \mathcal{C} \in \Pi, \tau \in \mathcal{H} \end{aligned} \quad (4)$$

where coefficient $\lambda_1 \geq 0$ controls the complexity of the mapping function in the ambient space while $\lambda_2 \geq 0$ controls the complexity of the mapping function in the intrinsic geometry. If $\lambda_2 = 0$, the LapGF becomes to the GF algorithm.

Let $\mathcal{K} : \mathbf{X} \times \mathbf{X} \mapsto \mathbb{R}^{d \times d}$ be a standard Mercer kernel with an associated RKHS family of functions $\mathcal{H}_{\mathcal{K}}$ with the corresponding norm $\|\cdot\|_{\mathcal{H}}$. We use $\|\tau\|_{\mathcal{M}}^2$ to measure the smoothness of τ . Then the optimal mapping function can be solved by minimizing the Eq. 3 under local Tikhonov and global manifold regularization.

More precisely, let \mathbf{W} be a nearest neighbor graph which serves as a discrete probe for the geometric structure of the data, then the graph Laplacian [7] is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$ which provides a natural intrinsic measure for simplicity of data-dependent smoothness,

$$\|\tau\|_{\mathcal{M}}^2 = \boldsymbol{\tau}^T \mathbf{L} \boldsymbol{\tau} = \frac{1}{2} \sum_{i,j} \mathbf{W}_{ij} \|\tau(x_i) - \tau(x_j)\|^2, \quad (5)$$

where $\boldsymbol{\tau} = (\tau(x_1), \dots, \tau(x_N))$, and \mathbf{D} is a diagonal matrix with elements $\mathbf{D}_{ii} = \sum_j \mathbf{W}_{ij}$. A conditional distribution τ is sufficiently smooth on the data manifold, and it is

need to ensure that if x_i is close to x_j , then $\tau(x_i)$ is close to $\tau(x_j)$ as well when minimizing the Laplacian regularization term,

2.4. Transformation Estimation

We choose a Gaussian kernel for \mathcal{K} with elements $k_{ij} = \exp(-\beta \|x_i - x_j\|^2)$, because it satisfies symmetric and positive define properties, and it makes the regularization terms easy to rewrite under the Representer theorem [1], then the estimated transformation τ by minimizing Eq. 4 takes the form of the Gaussian radial basis function

$$\tau(x_p) = \sum_{i=1}^N \alpha_i \mathcal{K}(x_p, x_i) = \mathcal{K}\mathbf{A}, \quad (6)$$

where the matrix $\mathcal{A}_{N \times d}$ with elements $(\alpha_1, \dots, \alpha_N)^T$ denotes the Gaussian kernel weights. Substituting Eq. 6 back into Eq. 4, and rewriting the objective function (Eq. 4) in matrix form,

$$\mathcal{G}(\mathbf{C}, \mathbf{A}) = \mathbf{E} + \lambda_1 \text{tr}(\mathbf{A}^T \mathcal{K} \mathbf{A}) + \lambda_2 \text{tr}(\mathbf{A}^T \mathcal{K}^T \mathbf{L} \mathcal{K} \mathbf{A}), \quad (7)$$

where $\mathbf{E} = \exp(-\|CY - X - \mathcal{K}\mathbf{A}\|^2/\sigma^2)$.

Then the non-rigid transformation can be obtained by the estimated optimal weight $\mathbf{A}^* = \text{argmin} \mathcal{G}(\mathbf{C}, \mathbf{A})$. Taking the derivative of Eq. 7 with respect to weight \mathbf{A} , due to the continuous differentiability of LapGF, we can obtain

$$\frac{\partial \mathcal{G}(\mathbf{C}, \mathbf{A})}{\partial \mathbf{A}} = \frac{2}{\sigma^2} \mathcal{K}^T (\mathbf{X} + \mathcal{K}\mathbf{A} - CY) \circ (\mathbf{E} \otimes \mathbf{1}) + 2\lambda_1 \mathcal{K}\mathbf{A} + 2\lambda_2 \mathcal{K}^T \mathbf{L} \mathcal{K} \mathbf{A} = 0. \quad (8)$$

In this paper, the objective is not convex, and it is unlikely that any algorithm can find its global minimum. However, a stable local minimum is often enough for many practical applications, due to the Gaussian fields is differentiable and preferably convex in the neighborhood of the optimal registered position. Thus, the numerical optimization problem can be solved by employing the gradient-based quasi-Newton method with deterministic annealing algorithm (see Section 2.6). As the iterations continue, we can get a good chance of reaching a stable local minimum.

2.5. Approximate Kernel Matrix

The kernel matrix plays an important role in the regularization theory, for instance, it provides an easy way to choose an RKHS. However, in this paper, the time complexity of Gaussian fields algorithm is approximately $\mathcal{O}(N^2M + N^3)$, and the performance will become poor as increasing the number of the points. Hopefully, low-rank kernel matrix approximation can yield a large increase in

speed with little loss in accuracy. As discussed in [20], the low-rank kernel matrix approximation constrains both the nonrigid transformation and its space. Choosing small rank of the matrix, the low-rank matrix approximation can be sufficient and accurate when facing a large number and well clustered points data. Precisely, low-rank kernel matrix approximation $\widehat{\mathcal{K}}$ is the closest N_l -rank matrix approximation to \mathcal{K} , and satisfies the Frobenius norm $\|\cdot\|_F$,

$$\underset{\widehat{\mathcal{K}}}{\text{argmin}} \|\mathcal{K} - \widehat{\mathcal{K}}\|_F, \quad s.t. \text{ rank}(\widehat{\mathcal{K}}) \leq N_l. \quad (9)$$

Applying the eigenvalue decomposition of \mathcal{K} , the approximated kernel matrix can be written as $\widehat{\mathcal{K}} = \mathbf{V} \Lambda \mathbf{V}^T$, where Λ is a diagonal matrix of size $N_l \times N_l$ with N_l largest eigenvalues and \mathbf{V} is an $N \times N_l$ matrix with the corresponding eigenvectors. Then the objective function (Eq. 7) and its derivative (Eq. 8) can be rewritten by substituting the low-rank kernel matrix $\widehat{\mathcal{K}}$,

$$\mathcal{G} = \widehat{\mathbf{E}} + \lambda_1 \text{tr}(\widehat{\mathbf{A}}^T \mathbf{P} \widehat{\mathbf{A}}) + \lambda_2 \text{tr}(\widehat{\mathbf{A}}^T \mathbf{Q} \widehat{\mathbf{A}}), \quad (10)$$

$$\begin{aligned} \frac{\partial \mathcal{G}}{\partial \widehat{\mathbf{A}}} = & \frac{2}{\sigma^2} \mathbf{P}^T (\mathbf{X} + \mathbf{P} \widehat{\mathbf{A}} - CY) \circ (\widehat{\mathbf{E}} \otimes \mathbf{1}) \\ & + 2\lambda_1 \mathbf{P} \widehat{\mathbf{A}} + 2\lambda_2 \mathbf{Q} \widehat{\mathbf{A}} = 0, \end{aligned} \quad (11)$$

where the newly weight parameter matrix $\widehat{\mathbf{A}}_{N_l \times d}$ with elements $(\alpha_1, \dots, \alpha_{N_l})^T$, $\Lambda_{N_l \times N_l}$ is a diagonal matrix, $\widehat{\mathbf{E}} = \exp(-\|CY - \mathbf{X} - \mathbf{P} \widehat{\mathbf{A}}\|^2/\sigma^2)$, $\mathbf{P}_{N \times N_l} = \mathbf{V} \Lambda$, and $\mathbf{Q}_{N_l \times N_l} = \mathbf{P}^T \mathbf{L} \mathbf{P}$.

By using the low-rank kernel matrix approximation, the time complexity is reduced down to $\mathcal{O}(NM)$ with $N_l \ll N$ approximately.

2.6. Implementation Details

In the optimization, we use a rigid to non-rigid strategy by applying deterministic annealing technique on the scale parameter σ^2 and β to improve the algorithm convergence. More specially, given a large initial value of σ^2 and β for global rigid transformation, and reducing them with a fixed annealing rate γ towards for local non-rigid transformation by equations $\sigma^2 = \gamma \sigma^2$, and $\beta = \gamma \beta$ iteratively. We empirically set $\sigma^2 = 2$, $\beta = 10$ and $\gamma = 0.93$ throughout this paper. Due to the optimization needs a termination condition, we choose a lower bound of σ^2 and set $\sigma_{final}^2 = 0.01$. Note that we also set a lower bound $\beta_{final} = 0.2$ to control the degree of non-rigid transformation, and β will be fixed when $\beta \leq \beta_{final}$. The experiments show that the method will catch convergence after about 30 iterations, as showed in Fig. 1. The parameters of regularization terms includes λ_1, λ_2 which are used to trade-off the smoothness, and the

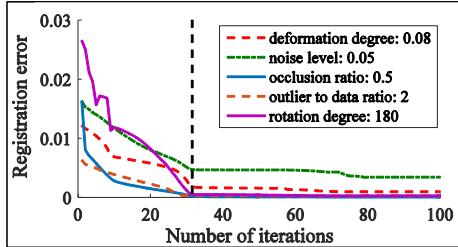


Figure 1. Convergence experiment on synthesized dataset under the largest degree of deformation, noise, occlusion, outlier, and rotation.

analysis of model selection is shown in Fig. 2. The test result shows that the proposed algorithm performs best for $\lambda_1 \in [0.1, 1]$ and $\lambda_2 \in [2, 8]$. In this paper, we fixed them as $\lambda_1 = 0.1$, $\lambda_2 = 5$. Note the construction of graph Laplacian, we choose heat kernel with band $\beta_s = 1$ to define the weight matrix \mathbf{W} , and the number of the nearest neighbor is set as $N_n = 3$. We use the Matlab implementation of the Laplacian regularization²[1] in our method.

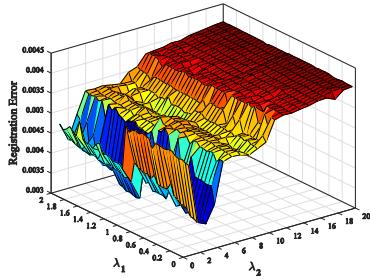


Figure 2. Model selection of the regularization parameters λ_1 and λ_2 for point set registration.

Low-rank kernel matrix approximation is applied to the method to reduce the computational complexity, the important parameter N_l denotes the number of the selected eigenvalues, and as a trade-off, it controls the balance between runtime complexity and registration accuracy. The proposed algorithm performs best for $N_l = [15, 20]$, where eigenvectors \mathbf{V} and eigenvalues Λ are computed by the fast Gauss transform (FGT) [9], and its experimental analysis is shown in Fig. 3.

All tested point sets are normalized to zero mean and unit variance by data normalization method (translation and re-scaling) for point set registration at the beginning of the experiments.

²manifold.cs.uchicago.edu/manifold_regularization/manifold.html

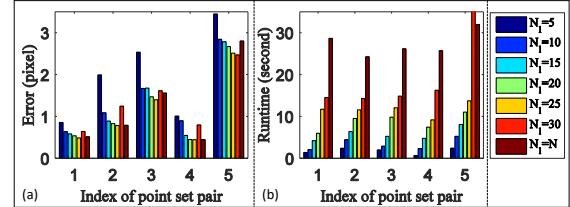


Figure 3. Registration results on IMM face landmarks under different values of N_l for low-rank kernel matrix approximation. (a) Statistics of registration errors. (b) Statistics of registration runtime.

3. Experiments

The proposed algorithm is implemented in Matlab, and tested on an Intel Core i5 CPU 2.5GHz with 8GB RAM.

3.1. Experimental Setup

Datasets. We use a variety of public datasets which are frequently used in the point set registration and robust point matching research. 1) **Synthesized Data.** This dataset³ is constructed by Chui and Rangarajan [6], and it consists of two different point sets: Chinese character and fish shape. 105 points are sampled from a Chinese character, and 98 points are sampled from the outmost silhouette of a fish. For each point set, five degradation categories, i.e., deformation, noise, occlusion, outliers, and rotation are used to evaluate the accuracy and robustness of PSR methods, and this dataset contains 5000 pairs of point set. 2) **IMM Face Database.** This database⁴ consists of the facial and multi-view changes. 58 point landmarks are sampled from a face with different facial expressions and poses. 3) **WILLOW Object Class Dataset.** This dataset⁵ contains five sets of real images with manually labeled ground-truth landmarks (10 points). 4) **Tools 2D Database.** Two-dimensional articulated shapes⁶ for non-rigid shape similarity experiments [5]. This dataset consists of 7 different articulated shapes. 5) **Oxford Affine Dataset.** This dataset⁷ consists of six different changes in imaging conditions: rotation, viewpoint changes, scale changes, image blur, illumination, and JPEG compression.

Comparisons. Non-rigid point set registration methods: TPS-RPM [6], SC [2], QPCCP [13], GMMReg [10], CPD [20], and RPM-L2E [18]. Mismatch removal methods: Random Sample Consensus (RANSAC) [8], Identify Correspondence Function (ICF) [12] based on the support vector regression, non-rigid RANSAC [21], and Vector Field Con-

³www.cise.ufl.edu/~anand/students/chui

⁴www.imm.dtu.dk/~aam/datasets/datasets.html

⁵www.di.ens.fr/willow/research/graphlearning

⁶tosca.cs.technion.ac.il/book/resources_data.html

html

⁷www.robots.ox.ac.uk/~vgg/data/data-aff.html

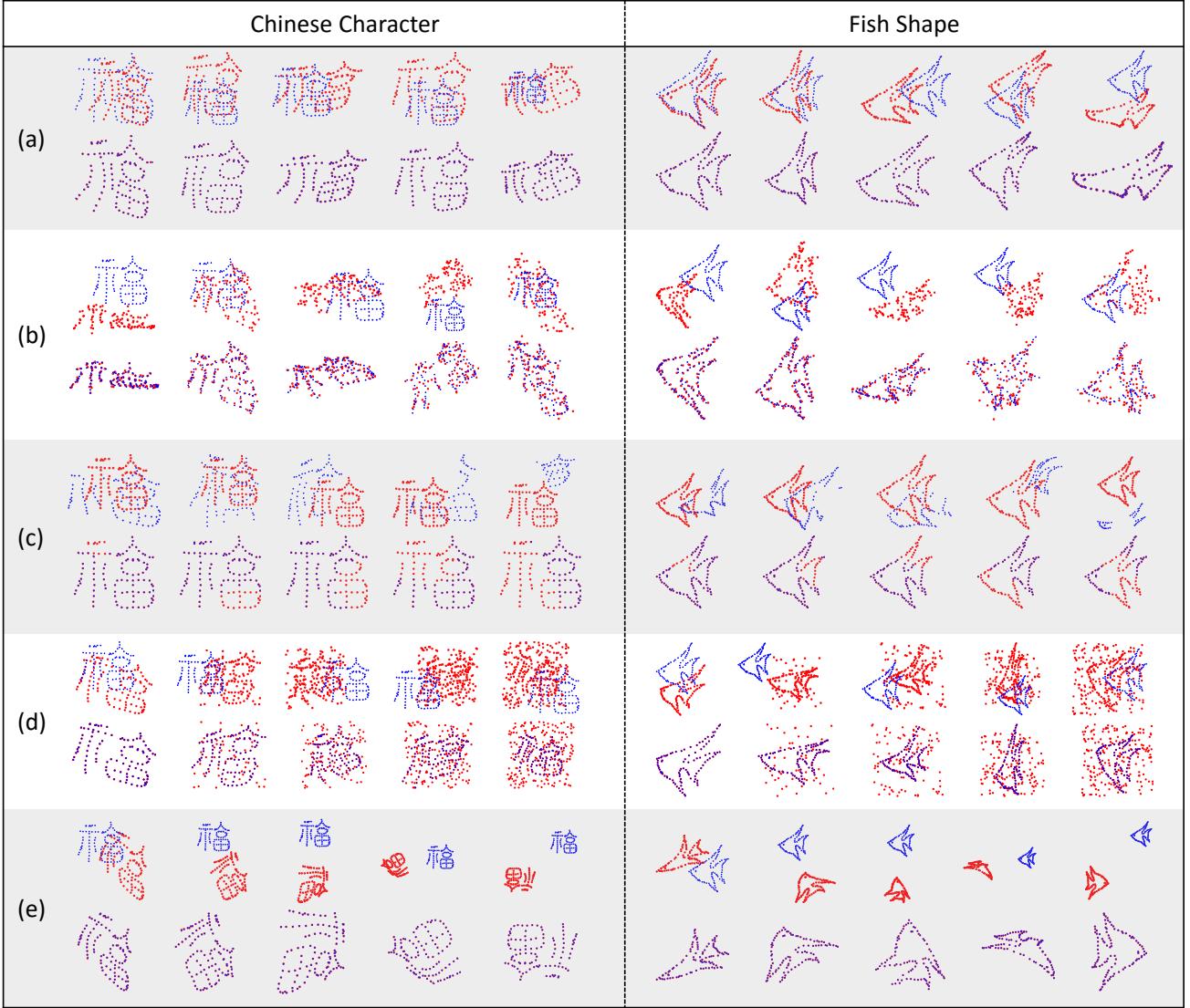


Figure 4. Registration results of the proposed CA-LapGF algorithm on the synthesized data: Chinese character and fish shape. (a) Deformation. (b) Noise. (c) Occlusion. (d) Outliers. (e) Rotation. From left to right in each group, the gradation level becomes larger.

sensus (VFC) [29, 19, 17]. All methods are implemented in Matlab, and tested on the same environment.

3.2. Results on Non-rigid PSR

3.2.1 Synthesized Data

Registration results of the proposed algorithm are shown in Fig. 4 on both Chinese character and fish shape point sets. In each degradation category, five degradation levels are designed to test the robustness of the PSR methods, where 100 point set pairs are created for each gradation level. The qualitative experimental results in the figure show that the model point sets (blue '+') are all well aligned onto the scene sets (red 'o') except the scene sets are distorted by

some degrees of noise (Fig. 4b), where the positions of the points in the scene set are disturbed by a certain degree of white Gaussian noise. It is worth noting that almost perfect registration results are shown under deformation, occlusion, outliers, and rotation degradations.

Fig. 5 shows the average registration error of several non-rigid PSR methods using the root-mean-square error (RMSE) on the synthesize data. Quantitative experimental comparison results demonstrate that the proposed CA-LapGF algorithm gets the lowest registration error on the whole tested scenarios. This is due to the scale, translation, and rotation invariant inner distance based context strategy can efficiently establish likely correspondences, and the CA-LapGF with global to local regularization refinement

helps to estimate non-rigid transformations robustly. Fig. 6 shows the average runtime of the algorithm with about 80 iterations, we can see that the runtime becomes larger under the noise and outlier degradations as the degree level increases. Formally, the algorithm takes about 15 seconds to align two point sets with $N, M = 100$ points.

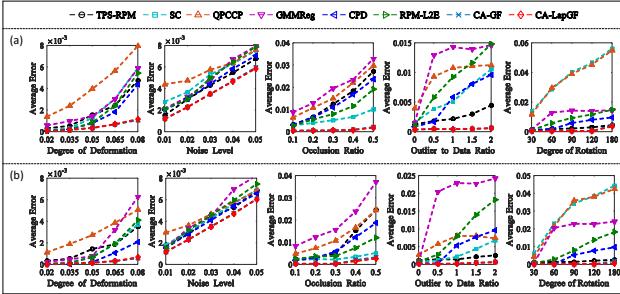


Figure 5. Comparison between six methods and our CA-GF and CA-LapGF to register point sets on Chinese character and fish shape. (a) Registration errors plot on Chinese character data. (a) Registration errors plot on fish shape.

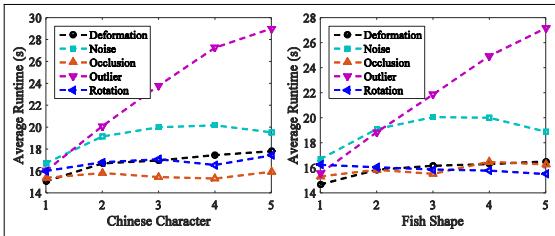


Figure 6. The average runtime of the registration on synthesized data. In each figure, the degree level becomes larger from 1 to 5.

3.2.2 IMM Face Database

In this experiment, we first choose a front face image, and let it be the model set, the other five images are defined as the scene set: from scene 1 to scene 5, as shown in Fig. 7a and Fig. 7b. The non-rigid deformation of the scene set becomes large as increasing the degree of viewpoint and facial expression. The qualitative registration results by aligning the model set onto the five scene sets respectively are shown in Fig. 7c. We can see that the point sets are well aligned by the CA-LapGF algorithm, however, the result on scene 5 is slightly bad due to the sampled landmarks are contaminated by noise when suffering from large posture and expression change. As shown in the rightmost figure of Fig. 7, the proposed CA-LapGF can get the better registration performance than the well-known SC [2], GMMReg [10], and CPD [20] methods on five groups of faces, where

the shape context registration method uses the TPS transformation model.

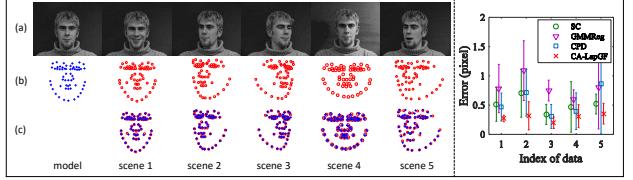


Figure 7. Registration results on face landmarks. (a) The IMM face images. (b) The face point landmarks. (c) The registration results of the CA-LapGF. The rightmost figure is the comparison between three PSR methods and the CA-LapGF, and the error bars indicate the registration error means and deviations over 5 samples in each group of data.

3.2.3 WILLOW Object Class Dataset

This real natural image dataset consists of five different object instances such as car, duck, face, motorbike, and winebottle, and we use this data to evaluate the performance for point matching. In order to match objects correctly, the model sets which are sampled from the object images are used to find their underlying correspondences by aligning onto the fixed scene set. An example of the experimental results on the datasets is shown in Fig. 8, where car, duck, and motorbike object images are selected in the experiment. The left group of figures show the multi-point set registration results on the given images which are shown in the right group of figures in Fig. 8. Perfectly registration results and the point matching are obtained by the CA-LapGF algorithm simultaneously, and it is easy to handle the viewpoint, posture, appearance and shape change. This is due to the proposed algorithm extracts the structure preserving inner distance and the rigid-to-non-rigid coarse-to-fine strategy.

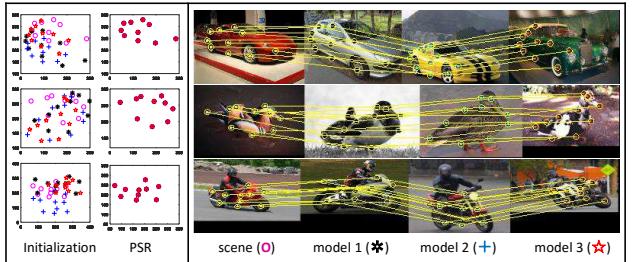


Figure 8. An example of point matching result on WILLOW object class dataset. Left: the initialization and registration results of multi-point sets, three model sets are aligned onto the fixed scene set. Right: the matching results of multi-feature point sets.

3.2.4 Tools 2D Database

This experiment tests the performance of the point set registration algorithms on articulated shapes. We compare CA-LapGF with three PSR algorithms: SC [2], GMMReg [10] and CPD [20] qualitatively. The initial point set contains 150 points which are sampled from the contour of each shape randomly, and the generated point sets without ground-truth are mainly used to test the performance of non-rigid transformation estimation. Fig. 9 shows the registration results of SC, GMMReg, CPD, and CA-LapGF algorithms. It can be observed that in the given five cases, CA-LapGF consistently achieves the best performance. The experimental results reveal that the advantages of the context-aware robust point set registration algorithm over other methods in solving general PSR problems.

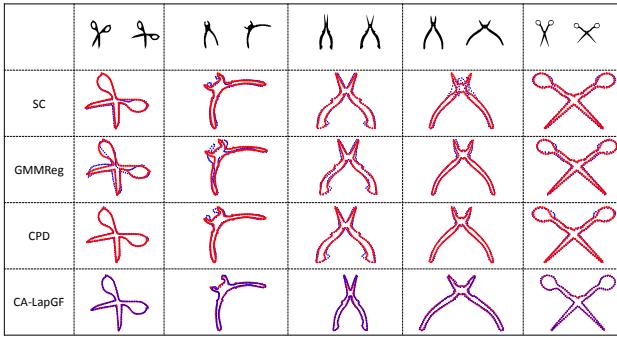


Figure 9. Examples of the registration results on articulated tools database. Five groups of real images are tested, and in each group, the left one is the model set, the right one is the scene set. Comparison between SC, GMMReg, CPD and CA-LapGF to align the articulated shapes.

3.3. Results on Mismatch Removal

The Oxford image dataset is used to test the mismatch removal performance of the proposed CA-LapGF algorithm with the known correspondences parameter C . We use the accuracy as the mismatch removal evaluation metric. We compare CA-LapGF with five robust point matching algorithms: RANSAC [8], ICF [12], CPD [20], Vector Field Consensus (VFC) [29], and non-rigid RANSAC [21]. It is worth noting that the correct matches are estimated by the Gaussian fields, $\mathcal{S} = \exp\left(-\frac{C\mathbf{Y}-\widehat{\mathbf{X}}}{\sigma^2}\right) \geq \xi$, where $\widehat{\mathbf{X}}$ is the transformed \mathbf{X} after several iterations, $\xi \in [0, 1]$ is a threshold which controls the precision and recall values, and we set $\xi = 0.1$ throughout this paper.

We use the VLFeat [23] toolbox with default settings to extract the feature points of each image, and the putative matches are generated by the nearest neighbour matching method (its threshold is 1.5). Input initial matches C_0 which contains some degrees of mismatches into the CA-

LapGF algorithm, and then output the index of the correct matches \mathcal{S} . We can evaluate the accuracy of the mismatch removal by comparing \mathcal{S} with the ground-truth data.

The quantitative comparison results are shown in Fig. 10. From image pairs '1v2' to '1v6', the transformation becomes larger, and the mismatch removal confronts with the challenge. The Comparison curves demonstrate that CA-LapGF achieves the better performance than other methods in most cases.

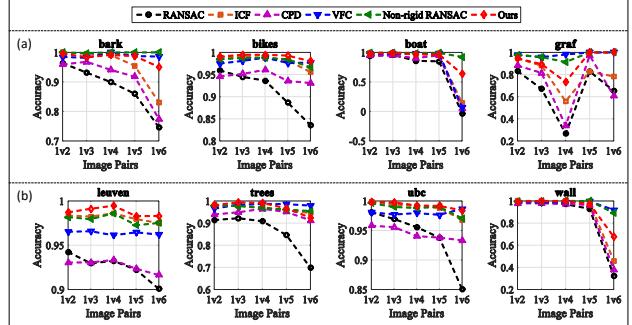


Figure 10. Comparison between mismatch removal methods and our CA-LapGF to remove mismatches for robust point matching.

4. Conclusion

This paper proposes CA-LapGF for non-rigid PSR and mismatch removal. The main idea for CA-LapGF is a novel robust non-rigid transformation estimation with inner distance based context-aware Gaussian fields. The biggest difference with density estimation methods [20, 26] is that our method uses an estimator based on Gaussian fields instead of building a more complex model that includes inliers and outliers. The benefits include: 1) inner distance based context descriptor captures the invariant feature and preserves the shape structure; 2) global and local regularization constrains the geometric transformation, and global rigid and local non-rigid coarse-to-fine technique makes the transformation estimation smooth; 3) our experiments demonstrate that much more accuracy can be achieved than the methods in state-of-the-art.

Acknowledgments

This work was partially supported by National Natural Science Foundation of China (NSFC 61103070), and the Fundamental Research Funds for The Central Universities.

References

- [1] M. Belkin, P. Niyogi, and V. Sindhwani. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *The Journal of Machine Learning Research*, 7:2399–2434, 2006.

- [2] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):509–522, 2002.
- [3] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
- [4] F. Boughezal, M. Mercimek, A. Koschan, and M. Abidi. A new method for the registration of three-dimensional point-sets: The gaussian fields framework. *Image and Vision Computing*, 28(1):124–137, 2010.
- [5] A. M. Bronstein, M. M. Bronstein, A. M. Bruckstein, and R. Kimmel. Analysis of two-dimensional non-rigid shapes. *International Journal of Computer Vision*, 78(1):67–88, 2008.
- [6] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2):114–141, 2003.
- [7] F. R. Chung. *Spectral graph theory*, volume 92. American Mathematical Soc., 1997.
- [8] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [9] L. Greengard and J. Strain. The fast gauss transform. *SIAM Journal on Scientific and Statistical Computing*, 12(1):79–94, 1991.
- [10] B. Jian and B. C. Vemuri. Robust point set registration using gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–1645, 2011.
- [11] H. Li, T. Shen, and X. Huang. Global optimization for alignment of generalized shapes. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 856–863. IEEE, 2009.
- [12] X. Li and Z. Hu. Rejecting mismatches by correspondence function. *International Journal of Computer Vision*, 89(1):1–17, 2010.
- [13] W. Lian, L. Zhang, Y. Liang, and Q. Pan. A quadratic programming based cluster correspondence projection algorithm for fast point matching. *Computer Vision and Image Understanding*, 114(3):322–333, 2010.
- [14] H. Ling and D. W. Jacobs. Shape classification using the inner-distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(2):286–299, 2007.
- [15] J. Ma, W. Qiu, J. Zhao, Y. Ma, A. L. Yuille, and Z. Tu. Robust l2e estimation of transformation for non-rigid registration. *IEEE Transactions on Signal Processing*, 63(5):1115–1129, 2015.
- [16] J. Ma, J. Zhao, Y. Ma, and J. Tian. Non-rigid visible and infrared face registration via regularized gaussian fields criterion. *Pattern Recognition*, 48(3):772–784, 2015.
- [17] J. Ma, J. Zhao, J. Tian, X. Bai, and Z. Tu. Regularized vector field learning with sparse approximation for mismatch removal. *Pattern Recognition*, 46(12):3519–3532, 2013.
- [18] J. Ma, J. Zhao, J. Tian, Z. Tu, and A. L. Yuille. Robust estimation of nonrigid transformation for point set registration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2147–2154. IEEE, 2013.
- [19] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu. Robust point matching via vector field consensus. *IEEE Transactions on Image Processing*, 23(4):1706 – 1721, 2014.
- [20] A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, 2010.
- [21] Q.-H. Tran, T.-J. Chin, G. Carneiro, M. S. Brown, and D. Suter. In defence of ransac for outlier rejection in deformable registration. In *Computer Vision-ECCV 2012*, pages 274–287. Springer, 2012.
- [22] Y. Tsin and T. Kanade. A correlation-based approach to robust point set registration. In *Computer Vision-ECCV 2004*, pages 558–569. Springer, 2004.
- [23] A. Vedaldi and B. Fulkerson. Vlfeat: An open and portable library of computer vision algorithms. In *Proceedings of the international conference on Multimedia*, pages 1469–1472. ACM, 2010.
- [24] G. Wang, Z. Wang, Y. Chen, and W. Zhao. A robust non-rigid point set registration method based on asymmetric gaussian representation. *Computer Vision and Image Understanding*, 141:67–80, 2015.
- [25] G. Wang, Z. Wang, Y. Chen, and W. Zhao. Robust point matching method for multimodal retinal image registration. *Biomedical Signal Processing and Control*, 19:68–76, 2015.
- [26] G. Wang, Z. Wang, Y. Chen, W. Zhao, and X. Liu. Fuzzy correspondences and kernel density estimation for contaminated point set registration. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1936–1941. IEEE, 2015.
- [27] G. Wang, Z. Wang, W. Zhao, and Q. Zhou. Robust point matching using mixture of asymmetric gaussians for non-rigid transformation. In *Computer Vision-ACCV 2014*, pages 433–444. Springer, 2015.
- [28] A. L. Yuille and N. M. Grzywacz. The motion coherence theory. In *IEEE 2nd International Conference on Computer Vision (ICCV)*, pages 344–353. IEEE, 1988.
- [29] J. Zhao, J. Ma, J. Tian, J. Ma, and D. Zhang. A robust method for vector field learning with application to mismatch removing. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2977–2984. IEEE, 2011.
- [30] Y. Zheng and D. Doermann. Robust point matching for nonrigid shapes by preserving local neighborhood structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):643–649, 2006.