

# Robust Non-Rigid Point Set Registration Using Spatially Constrained Gaussian Fields

Gang Wang, Qiangqiang Zhou, and Yufei Chen

**Abstract**—Estimating transformations from degraded point sets is necessary for many computer vision and pattern recognition applications. In this paper, we propose a robust non-rigid point set registration method based on spatially constrained context-aware Gaussian fields. We first construct a context-aware representation (e.g., shape context) for assignment initialization. Then, we use a graph Laplacian regularized Gaussian fields to estimate the underlying transformation from the likely correspondences. On the one hand, the intrinsic manifold is considered and used to preserve the geometrical structure, and *a priori* knowledge of the point set is extracted. On the other hand, by using the deterministic annealing, the presented method is extended to a projected high-dimensional feature space, i.e., reproducing kernel Hilbert space through a kernel trick to solve the transformation, in which the local structure is propagated by the coarse-to-fine scaling strategy. In this way, the proposed method gradually recovers much more correct correspondences, and then estimates the transformation parameters accurately and robustly when facing degradations. Experimental results on 2D and 3D synthetic and real data (point sets) demonstrate that the proposed method reaches better performance than the state-of-the-art algorithms.

**Index Terms**—Gaussian fields, registration, non-rigid transformation, graph Laplacian regularization, correspondence.

## I. INTRODUCTION

THE problem of estimating underlying transformations and recovering correspondences in point sets from degradations (deformation, noise, occlusion, outliers, rotation, multi-view changes, etc.) is intractable in many computer vision, pattern recognition, and medical imaging applications.

Generally, point set registration is mainly categorized into rigid and non-rigid when considering the transformation pattern. We are particularly interested in the non-rigid point set registration because the underlying non-rigid transformation model is usually unknown and challenging to define or approximate. As a key component in point set registration, the non-rigid transformation estimation exists in many engineering

Manuscript received September 10, 2016; revised December 19, 2016 and January 16, 2017; accepted January 23, 2017. Date of publication January 24, 2017; date of current version February 17, 2017. This work was supported by the National Natural Science Foundation of China under Grant 61573235. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Vladimir Stankovic.

G. Wang is with the School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China (e-mail: gwang.cv@gmail.com).

Q. Zhou and Y. Chen are with the College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China (e-mail: zqqcsu@gmail.com; yufechen@tongji.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2017.2658947

and science applications, including hand-written character recognition, fingerprint identification, facial-expression recognition, and non-rigid medical image registration.

However, in many problems of point set registration, extracted feature point sets from images or point cloud from real world scenes often contain outliers that should be rejected for improving the accuracy of the registration task. This correspondence problem is one of the primary challenges. Solving for non-rigid transformations between the model and scene point sets with contaminated correspondences is another challenge. Precisely, the correspondences recovering accuracy is sensitive to the data degradations which make the distribution of a point set become more complex. Moreover, the transformation estimation accuracy is sensitive to the local minima in numerical optimization. With these challenges, point set registration problem becomes increasingly difficult. Typically, estimating non-rigid transformations from putative correspondences (which are recovered from degraded point sets) is inherently an ill-posed inverse problem. The performance of a point set registration algorithm mainly depends on how well it can employ spatially constrained conditions or priors when numerically solving the problem because the useful prior statistical knowledge can regulate estimated non-rigid transformation parameters.

In this paper, we focus on the non-rigid transformation and introduce a robust non-rigid transformation estimation algorithm: spatially constrained Gaussian fields (SCGF) for point set registration. The proposed SCGF tries to address the registration problem accurately and robustly in the presence of the aforementioned degradations and to overcome the limitations of the existing algorithms. Briefly, the key ideas of our method are (a) to find meaningful correspondences using the context information of points by computing their histograms, and (b) to estimate the underlying non-rigid transformation using robust point matching by graph Laplacian regularized Gaussian fields. More precisely, We try to extract low dimensional geometrical structure and use it as prior knowledge to constrain the process of point set registration. In particular, we add as a prior the intrinsic manifold structure by making use of point set correspondences for modeling the algorithm. The estimated correspondences are designed as the attribute weight of the Gaussian fields and a robust estimation of non-rigid transformation can be obtained. Motivated by the properties of reproducing kernel Hilbert spaces (RKHS) with the Representer theorem [1], non-rigid transformations can be mapped into the constructed 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 graph Laplacian regularization term is added into the objective function. Under the determined annealing framework, the objective function can be optimized by the quasi-Newton technique. At the same time, the availability of multi-scale structures can significantly make the ill-posed problem to be better posed. Moreover, low-rank kernel matrix approximation is applied to reduce runtime when facing a large number of points. Extensive experiments on some synthesized and real image datasets demonstrate that SCGF is more robust in the presence of a large degree of degradations (deformation, noise, occlusion, outlier, rotation and multi-view changes).

The rest of the paper is organized as follows: In Section II, we briefly overview the current methods and state our contributions. In Section III, we first introduce the related work of point set registration via Gaussian fields framework, then we present the proposed spatially constrained Gaussian fields based on graph Laplacian regularization. Section IV details the algorithm implementation. Section V presents the experimental setup, results, and comparative studies. Section VI briefly concludes the paper.

## II. PREVIOUS WORK

To overcome the aforementioned problems and challenges, a vast variety of point set registration algorithms have been proposed in the literature. Here, we overview them and state our contributions briefly.

The Iterative Closest Point (ICP) algorithm [2] has simplicity and low computational complexity, which uses the nearest-neighbor distance criterion to assign binary correspondences (closest point computation), and the least squares to estimate the rigid transformation iteratively. However, ICP requires an initial position, in fact, if it gets true correspondences, and then the estimation procedure could operate in one shot. Typically, an additional regularization term is adding to improve ICP for non-rigid transformation point set registration. For non-rigid transformation, Chui and Rangarajan [3] introduced a soft assignment technique and 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 has high computational complexity. Zheng et al. [4] proposed a robust point matching by preserving local neighborhood structures (RPM-PLNS) for non-rigid shape registration. They use a graph matching technique to preserve local neighborhood structures, but their algorithm is still sensitive to outliers. Kernel correlation (KC) [5] 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. Subsequently, based on the theory of the KC, a robust point set registration approach using Gaussian mixture models (GMM) [6] has been presented (namely GMMReg), it leverages the closed-form expression for the  $L_2$  distance

between two Gaussian mixtures which are used to represent the given point sets.

From the view of motion coherence theory (MCT) [7], Myronenko et al. [8] 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, and the other is considered as the data point set. Then the correspondence problem is formulated as a density estimation problem. More precisely, Expectation-Maximization (EM) 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. Based on this approach, Ma et al. presented a robust method by preserving global and local structures [10].

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. They then apply particle swarm optimization (PSO) to recover the optimal transformation parameters instead of the gradient-based algorithms. Ma et al. [12], [13] 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 (namely RPM- $L_2$ E), where the algorithm needs putative correspondences estimated by 2D shape context descriptor for non-rigid point set registration. A non-rigid point set registration method based on asymmetric Gaussian representation [14], [15] uses a mixture of asymmetric Gaussians to represent point sets, and it updates correspondences and transformations under the framework of TPS-RPM. In addition, an interesting point matching method is presented using cluster correspondence projection with a quadratic programming (QPCCP) [16]. Although QPCCP is fast to compute the registration problem, it is sensitive to degradations.

Our proposed Gaussian field-based algorithm shares similarities to the existing Gaussian fields framework presented for three-dimensional rigid surface registration [17], and to its application non-rigid visible and infrared face images which is based on the Tikhonov regularization theory [18]. However, our spatially constrained context-aware Gaussian fields mainly focuses on the non-rigid transformation, and is applied to register non-rigid point sets and remove mismatches from a putative correspondence with more accuracy and robustness. Part of our previous work has been reported in [19]. Here, we mainly extend this work from two-dimension to three-dimension data (point set), and solve the objective function with GRBF or TPS kernel, and add more implementation details.

Briefly, the main contributions of our study includes: 1) an alternating iterative updating strategy between correspondences and transformations that lets us simplify the Gaussian mixture models to the spatially constrained Gaussian fields which can estimate the non-rigid transformations robustly; 2) a graph Laplacian regularization is formulated for

non-rigid point set registration by preserving low-dimensional manifold; 3) we apply the robust SCGF to two and three-dimensional non-rigid point set registration and correspondence rejecting, and the experimental results demonstrate that the proposed SCGF outperforms other state-of-the-art methods.

### III. POINT SET REGISTRATION VIA GAUSSIAN FIELDS

#### A. Problem Formulation

Given two point sets (see notation),<sup>1</sup> the model set  $X = \{x_i | x_i \in \mathbb{R}^d, i = 1, \dots, N\}$  and the scene set  $Y = \{y_j | y_j \in \mathbb{R}^d, j = 1, \dots, M\}$ , the scene set  $Y$  is fixed as a target set, and the model set  $X$  is moved onto the scene set by an optimize displacement transformation  $\mathcal{T}^*$ . We are assuming that there exists a significant overlap between the model and scene point sets. Given a shape attribute  $\mathcal{S}$ , used to guide the point set correspondence, the point set registration can be formulated as follows:

$$E(\mathcal{T}) = \lambda\phi(\mathcal{T}) + \sum_{i=1}^N \sum_{j=1}^M \exp \left[ -\frac{\|y_j - x'_i\|^2}{\sigma^2} - \frac{(\mathcal{S}(y_j) - \mathcal{S}(x'_i))^T \Sigma^{-1} (\mathcal{S}(y_j) - \mathcal{S}(x'_i))}{\xi^2} \right], \quad (1)$$

where  $x'_i = x_i + \mathcal{T}(x_i)$  denotes the transformed point,  $\xi$  and  $\sigma$  are scales,  $\Sigma$  denotes the covariance matrix of the feature descriptor,  $\phi(\cdot)$  is the Tikhonov regularization,  $\lambda \geq 0$  to control the level of smoothness. The Gaussian fields framework [17] uses  $\lambda = 0$  and an invariant moment attribute for  $\mathcal{S}$  to perform three-dimensional rigid pint set with transformation  $\mathcal{T}(x_i) = Rx_i + t$  ( $R$  denotes rotation,  $t$  denotes translation). Setting  $\lambda > 0$  and using two-dimensional shape context as  $\mathcal{S}$  for non-rigid transformation  $\mathcal{T}$  in [18]. It's worth noting that Gaussian fields is different from the Boolean registration criterion:

$$E(\mathcal{T}) = \sum_{i=1}^N \sum_{j=1}^M \delta[dist(x_i + \mathcal{T}(x_i), y_j)],$$

with  $\delta[dist(\cdot)] = \begin{cases} 1, & \text{for } dist(\cdot) = 0 \\ 0, & \text{otherwise,} \end{cases}$  (2)

where  $dist(\cdot)$  denotes the Euclidean distance between points. By using the Gaussian fields, we can convert the Boolean registration criterion into a smooth sum of Gaussians based on the mollification method [20].

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

#### B. Recovering Correspondence Weight via Context-Aware Descriptor

In Gaussian fields, we can use a correspondence weight matrix  $C$  to denote the attribute matching where  $C(x'_i, y_j) = c_{ij}$ .

In order to estimate the underlying Gaussian components of the Gaussian fields, a context-aware strategy is introduced to find the correspondence between points. For our registration problem, the assignment weight matrix  $C$  is a one-to-one mapping hard-assignment with  $\{0, 1\}$ .

$$c_{ij} = \begin{cases} 1, & \text{for } x'_i \rightarrow y_j, \text{ match} \\ 0, & \text{for } x'_i \not\rightarrow y_j, \text{ mismatch,} \end{cases} \quad (3)$$

where  $C$  can be estimated by the inner distance based context descriptor (IDSC) [21] and the three-dimensional context descriptor (3DSC) [22] for 2D and 3D point set registration, respectively.

More precisely, instead of using the attribute matching  $\exp[-(\mathcal{S}(y_j) - \mathcal{S}(x'_i))^T \Sigma^{-1} (\mathcal{S}(y_j) - \mathcal{S}(x'_i))/\xi^2]$ , the Chi-squared test and bipartite graph matching [23] are used to recover the correspondence weight. From the IDSC and 3DSC, we can extract the context information at point  $x_i$  by computing a histogram  $S_{X,i}$  of the relative coordinates of the remaining  $N - 1$  points  $\bar{x}_i$ :

$$S_{X,i}(k) = \#\{\bar{x}_i \neq x_i : dist(\bar{x}_i, x_i) \in \text{bin}(k)\}, \quad (4)$$

where  $dist(\cdot)$  is Euclidean and inner distance for 3DSC and IDSC, respectively. Note that the inner distance represents the shortest path distance between points within the point set outermost silhouette [21].

Then the matching  $\pi$  from  $X$  to  $Y$  can be assigned by minimizing (resorting to dynamical programming [21] or the Hungarian method [24]) the following total match cost:

$$H(\pi) = \sum_{i=1}^N \Omega(x_i, y_{\pi(i)}), \quad (5)$$

where the Chi-squared test statistic:

$$\Omega(x'_i, y_j) = \omega_{ij} = \frac{1}{2} \sum_{k=1}^K \frac{[S_{X,i}(k) - S_{Y,j}(k)]^2}{S_{X,i}(k) + S_{Y,j}(k)}, \quad (6)$$

and  $\pi$  is a permutation.

With the permutation result  $\pi(i)$ , we can recover the correspondence weight:

$$C(x'_i, y_j) = c_{ij} = \begin{cases} 1, & \text{if } \pi(i) \neq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Then the Gaussian fields proceeds by optimizing between  $C$  and  $\mathcal{T}$  in an alternating iterative method,

$$\begin{aligned} \underset{\mathcal{C}, \mathcal{T}}{\text{argmin}} \quad E(\mathcal{C}, \mathcal{T}) &= \lambda\phi(\mathcal{T}) + \sum_{i=1}^N \sum_{j=1}^M C(x'_i, y_j) \\ &\times \exp \left[ -\frac{\|y_j - (x_i + \mathcal{T}(x_i))\|^2}{\sigma^2} \right], \\ \text{s.t. } \mathcal{C} &\in \Pi, \quad \mathcal{T} \in \mathcal{H} \end{aligned} \quad (8)$$

where  $\Pi$  denotes a permutation matrix in each iteration, and  $\mathcal{H}$  is a reproducing kernel Hilbert space (RKHS).

### C. Gaussian Fields With Spatially Constraint

Motivated by the manifold assumption [25], we model the point set as a undirected graph, where the vertices are the points and the edges denote the relationships between vertices. In this way, we can resort to the graph Laplacian to describe the intrinsic geometrical structure of the data space. Mathematically, the manifold assumption can be implemented by minimizing the following term:

$$\mathcal{R}(\mathcal{T}) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N W_{ij} \|\mathcal{T}(x_i) - \mathcal{T}(x_j)\|^2, \quad (9)$$

where  $W_{ij}$  denotes the weight in a nearest neighbor graph which serves as the affinity between two vertices. Precisely, if points  $x_i$  and  $x_j$  are connected by an edge, there are three common ways to construct the weight matrix  $\mathbf{W}$ :

$$W_{ij} = \begin{cases} 1, & \text{0-1 weighting} \\ \exp(-\frac{\|x_i - x_j\|^2}{\sigma_s^2}), & \text{heat kernel} \\ x_i^T x_j, & \text{dot-product,} \end{cases} \quad (10)$$

where heat kernel weighting considers the geometrical nearby.

Define a diagonal matrix  $\mathbf{D}$  with elements  $D_{ii} = \sum_{j=1}^N W_{ij}$ , and then the graph Laplacian [26] is defined as  $\mathbf{L} = \mathbf{D} - \mathbf{W}$  which provides a natural intrinsic measure for simplicity of data-dependent smoothness. By spectral graph theory, Eq. (9) can be approximated as:

$$\begin{aligned} \mathcal{R}(\mathcal{T}) &= \sum_{i=1}^N \mathcal{T}(x_i)^2 D_{ii} - \sum_{i=1}^N \sum_{j=1}^N W_{ij} \mathcal{T}(x_i) \mathcal{T}(x_j) \\ &= \mathcal{T}^T \mathbf{D} \mathcal{T} - \mathcal{T}^T \mathbf{W} \mathcal{T} = \mathcal{T}^T \mathbf{L} \mathcal{T}, \end{aligned} \quad (11)$$

where  $\mathcal{T} = [\mathcal{T}(x_1), \dots, \mathcal{T}(x_N)]^T$ . For a conditional distribution  $\mathcal{T}$  to be sufficiently smooth on the data manifold, it needs to ensure that if  $x_i$  is close to  $x_j$ . Then  $\mathcal{T}(x_i)$  is close to  $\mathcal{T}(x_j)$  as well when minimizing the Laplacian regularization.

Now we can give the objective function of the spatially constrained Gaussian fields as follows:

$$\begin{aligned} E(\mathbf{C}, \mathcal{T}) &= \gamma \mathcal{R}(\mathcal{T}) + \lambda \phi(\mathcal{T}) + \sum_{i=1}^N \sum_{j=1}^M \mathbf{C}(x'_i, y_j) \\ &\quad \times \exp\left[-\frac{\|y_j - (x_i + \mathcal{T}(x_i))\|^2}{\sigma^2}\right], \end{aligned} \quad (12)$$

where  $\mathcal{R}(\mathcal{T}) = \|\mathcal{T}\|_{\mathcal{M}}^2$  is used to measure the smoothness of  $\mathcal{T}$ ,  $\phi(\mathcal{T}) = \|\mathcal{T}\|_{\mathcal{H}}^2$ ,  $\lambda > 0$  and  $\gamma > 0$  are the regularization parameters which control the complexity of the mapping function in the ambient space and intrinsic geometry, respectively. Then the optimal transformation function can be solved by minimizing the Eq. (12) under the Tikhonov and manifold regularization.

### D. Estimating Non-Rigid Transformation by Kernel

In order to estimate the non-rigid transformation and obtain the optimal solution of the objective function, 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  $\|\cdot\|_{\mathcal{H}}$ . Then a useful property of RKHS is shown as follows:

*Theorem 1 (Representer Theorem):* The minimization of the objective function (12) has a unique solution, given by

$$\mathcal{T}(x_p) = \sum_{i=1}^N \alpha_i \mathcal{K}(x_i, x_p). \quad (13)$$

Motivated by the above Representer Theorem [1], we can obtain the expression of transformation:

$$\mathcal{T} = \mathcal{K}\boldsymbol{\alpha}, \quad (14)$$

where  $\mathcal{K}$  is the kernel-valued matrix with elements  $K_{ij} = \mathcal{K}(x_i, x_j)$ , and  $\boldsymbol{\alpha}$  is an unknown coefficient vector  $[\alpha_1, \dots, \alpha_N]^T$  which denotes the kernel weights.

For the kernel selection, we use the radial basis function (RBF) to construct the kernel-valued matrix, including Gaussian RBF (GRBF) and thin-plate spline (TPS). More precisely, GRBF is defined as  $K_{ij} = \exp(-\beta \|x_i - x_j\|^2)$ , and TPS is defined as  $K_{ij} = r \log r = |x_i - x_j| \log |x_i - x_j|$  and  $K_{ij} = -r = -|x_i - x_j|$  for two-dimension and three-dimension, respectively, where  $|x_i - x_j|$  denotes the Euclidean distance.

1) *GRBF*: With the above defined GRBF kernel, substituting Eq. (14) back into Eq. (12), and rewriting the objective function in matrix form:

$$\begin{aligned} E(\boldsymbol{\alpha}) &= \gamma \text{trace}(\boldsymbol{\alpha}^T \mathcal{K}^T \mathbf{L} \mathcal{K} \boldsymbol{\alpha}) + \lambda \text{trace}(\boldsymbol{\alpha}^T \mathcal{K} \boldsymbol{\alpha}) \\ &\quad + \sum_{i=1}^N \sum_{j=1}^M C_{ij} \exp\left[-\frac{\|Y_j - (X + \mathcal{K}\boldsymbol{\alpha})_i\|^2}{\sigma^2}\right], \end{aligned} \quad (15)$$

then the non-rigid transformation can be obtained by the estimated optimal kernel weight  $\boldsymbol{\alpha}^* = \arg\min E(\boldsymbol{\alpha})$ . Hence, by the quasi-Newton method, the derivative of the final objective function  $E$  with respect to the GRBF parameter  $\boldsymbol{\alpha}$  can be obtained as:

$$\frac{\partial E(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = 2\gamma \mathcal{K}^T \mathbf{L} \mathcal{K} \boldsymbol{\alpha} + 2\lambda \mathcal{K} \boldsymbol{\alpha} - \mathcal{K}^T \frac{\partial E(\boldsymbol{\alpha})}{\partial \mathcal{T}}, \quad (16)$$

where if correspondence weight matrix achieves the optimal assignment  $\widehat{\mathbf{C}}$ , put

$$\frac{\partial E(\boldsymbol{\alpha})}{\partial \mathcal{T}} = -\frac{2}{\sigma^2} \mathbf{T}_{grbf} \circ \left[ \exp\left(-\frac{\|\mathbf{T}_{grbf}\|^2}{\sigma^2}\right) \otimes \mathbf{1} \right], \quad (17)$$

where  $\mathbf{T}_{grbf} = \widehat{\mathbf{Y}} - (X + \mathcal{K}\boldsymbol{\alpha})$ , and the weighted scene point set  $\widehat{\mathbf{Y}}$  is sorted according to  $\widehat{\mathbf{C}}$ .

2) *TPS*: Since the TPS transformation contains the affine part  $\mathbf{A}$ , and it always needs some control points  $\mathbf{J} = \{J_i | J_i \in \mathbb{R}^d, i = 1, \dots, k\}$  which can be selected from the model set randomly, then the TPS kernel ( $k$ -by- $k$ ) can be expressed as  $\mathcal{K}_{tps}(r)$ , where  $r = |J_i - J_j|$ . Note that this kernel matrix describes the internal structure of the control point set. Similarity, the basis matrix  $\mathbf{B}_{N \times k}$  can be obtained as  $B_{ij} = \mathcal{K}_{tps}(|x_i - J_j|)$ . Hence, the non-rigid transformation can be defined as  $\mathcal{T}_{tps} = [X | \mathbf{1}] \mathbf{A}^T + \mathbf{B} \boldsymbol{\tau}$ , where  $\boldsymbol{\tau}_{k \times d} = [\mathbf{A}^T; \mathbf{v}]$  is the TPS warping coefficient,  $\mathbf{A}_{d \times (d+1)} = [a_{11}, a_{12}, t_1; a_{21}, a_{22}, t_2]$ ,  $\mathbf{1}_{N \times 1}$  is a column vector. Note that  $\boldsymbol{\tau}^T [\mathbf{1} | \mathbf{J}] = \mathbf{0}$  ensures the TPS bending energy greater than zero and the nonlinear part of the transformation is zero at

infinity [6], so we can obtain  $\boldsymbol{\tau} = \mathbf{Z}\mathbf{v}$  where  $\mathbf{Z}_{k \times (k-d-1)}$  is the left null space of  $[\mathbf{1}|\mathbf{J}]$ ,  $\mathbf{v}$  is a  $(k-d-1) \times d$  matrix. Now we can rewrite the objective function as follows:

$$\begin{aligned} E(\mathbf{A}, \mathbf{v}) &= \gamma \operatorname{trace}(\mathbf{v}^T \mathbf{Z}^T \mathcal{K}_{tps}^T \mathbf{L} \mathcal{K}_{tps} \mathbf{Z} \mathbf{v}) \\ &\quad + \lambda \operatorname{trace}(\mathbf{v}^T \mathbf{Z}^T \mathcal{K}_{tps} \mathbf{Z} \mathbf{v}) + E_{tps}, \end{aligned} \quad (18)$$

where  $E_{tps} = \exp(-\|\widehat{\mathbf{Y}} - (\mathbf{X} + \mathcal{T}_{tps})\|^2/\sigma^2)$ .

Similarly, the derivatives of the final objective function  $E$  with respect to  $\mathbf{A}$  and  $\mathbf{v}$  can be obtained as:

$$\frac{\partial E(\mathbf{A}, \mathbf{v})}{\partial \mathbf{A}} = \frac{2}{\sigma^2} [\mathbf{X}|\mathbf{1}]^T \left[ \mathbf{T}_{tps} \circ \left( \exp\left(-\frac{\|\mathbf{T}_{tps}\|^2}{\sigma^2}\right) \otimes \mathbf{1} \right) \right], \quad (19)$$

$$\begin{aligned} \frac{\partial E(\mathbf{A}, \mathbf{v})}{\partial \mathbf{v}} &= 2\gamma \mathbf{Z}^T \mathcal{K}_{tps}^T \mathbf{L} \mathcal{K}_{tps} \mathbf{Z} \mathbf{v} + 2\lambda \mathbf{Z}^T \mathcal{K}_{tps} \mathbf{Z} \mathbf{v} \\ &\quad + \frac{2}{\sigma^2} \mathbf{Z}^T \mathbf{B}^T \left[ \mathbf{T}_{tps} \circ \left( \exp\left(-\frac{\|\mathbf{T}_{tps}\|^2}{\sigma^2}\right) \otimes \mathbf{1} \right) \right], \end{aligned} \quad (20)$$

where  $[\mathbf{X}|\mathbf{1}]^T$  is the homogeneous coordinates of the model point set, and  $\mathbf{T}_{tps} = \widehat{\mathbf{Y}} - (\mathbf{X} + \mathcal{T}_{tps})$ .

It is worth nothing that the objective function 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, and Gaussian fields provides this by being differentiable and preferably convex in the neighborhood of the optimally registered position. Thus, the numerical optimization problem can be solved by employing the gradient-based quasi-Newton method with deterministic annealing strategy (see Section IV). As the iterations proceed, the algorithm can reach a stable local minimum.

#### E. Kernel-Valued Matrix Approximation

The kernel-valued matrix plays a key role in the regularization theory, for instance, it provides an easy way to choose an RKHS. From the above transformation estimation with TPS kernel, selecting some control points can reduce the cost of computing kernel. Instead of using random selection, we can resort to making use of low-rank kernel-valued matrix approximation to speed up the computing for GRBF kernel. As discussed in the literature [8], the low-rank kernel-valued matrix approximation is used to constrain both the non-rigid transformation and its space. Choosing small rank of the matrix, the low-rank approximation can be sufficient and accurate when facing a large number of well-clustered points data.

More Precisely, low-rank kernel-valued matrix approximation  $\widehat{\mathcal{K}}$  is the closest  $n_{ctrl}$ -rank matrix approximation to  $\mathcal{K}$ , and satisfies the Frobenius norm  $\|\cdot\|_F$ ,

$$\begin{aligned} \arg \min_{\widehat{\mathcal{K}}} \|\mathcal{K} - \widehat{\mathcal{K}}\|_F, \\ \text{s.t. } \operatorname{rank}(\widehat{\mathcal{K}}) \leq n_{ctrl}. \end{aligned} \quad (21)$$

By applying the eigenvalue decomposition of  $\mathcal{K}$ , the approximated kernel-valued matrix can be written as:

$$\widehat{\mathcal{K}} = \mathbf{V} \Lambda \mathbf{V}^T \quad (22)$$

---

**Algorithm 1** Spatially Constrained Gaussian Fields (SCGF) for Non-Rigid Transformation Estimation

---

```

Input: The putative correspondences  $\widehat{\mathbf{C}}$ .
Output: The transformed parameter  $\alpha^*$ .
Initialize:  $\sigma, \beta, \lambda, \gamma, n_{ctrl}, \alpha = 0$ ;
Begin
Get  $\widehat{\mathbf{Y}}$  by resorting  $\mathbf{Y}$  according to  $\widehat{\mathbf{C}}$ ;
switch RBF type do
| case GRBF
|   Get the approximated matrix  $\widehat{\mathcal{K}}$ , and  $\mathcal{P}$ ;
| case TPS
|   Get the kernel matrix  $\mathcal{K}_{tps}$ , and basis  $\mathbf{B}$ ;
end
while optimization do
| Compute the Laplacian term by Eq. (11);
| Compute  $\mathbf{T} = \widehat{\mathbf{Y}} - (\mathbf{X} + \mathcal{T}(\mathbf{X}))$ ;
| Compute the energy  $E$  by Eq. (23, 18);
| Compute the gradient by Eq. (24, 19, 20);
| Update the parameter  $\alpha$ ;
end
return the optimal parameter  $\alpha^*$ ;

```

---

where  $\Lambda$  is a diagonal matrix of size  $n_{ctrl} \times n_{ctrl}$  with  $n_{ctrl}$  largest eigenvalues and  $\mathbf{V}$  is an  $N \times n_{ctrl}$  matrix with the corresponding eigenvectors.

Substituting the low-rank approximated kernel-valued matrix  $\widehat{\mathcal{K}}$ , then the objective function (15) and its derivative (16) can be rewritten as:

$$\begin{aligned} E(\widehat{\alpha}) &= \gamma \operatorname{trace}(\widehat{\alpha}^T \mathcal{P}^T \mathbf{L} \mathcal{P} \widehat{\alpha}) + \lambda \operatorname{trace}(\widehat{\alpha}^T \mathcal{P} \widehat{\alpha}) \\ &\quad + \sum_{i=1}^N \sum_{j=1}^M C_{ij} \exp\left[-\frac{\|\mathbf{Y}_j - (\mathbf{X} + \mathcal{P}\widehat{\alpha})_i\|^2}{\sigma^2}\right], \end{aligned} \quad (23)$$

$$\begin{aligned} \frac{\partial E(\widehat{\alpha})}{\partial \widehat{\alpha}} &= 2\gamma \mathcal{P}^T \mathbf{L} \mathcal{P} \widehat{\alpha} + 2\lambda \mathcal{P} \widehat{\alpha} + \frac{2}{\sigma^2} \mathcal{P}^T (\widehat{\mathbf{Y}} - (\mathbf{X} + \mathcal{P}\widehat{\alpha})) \\ &\quad \circ \left[ \exp\left(-\frac{\|\widehat{\mathbf{Y}} - (\mathbf{X} + \mathcal{P}\widehat{\alpha})\|^2}{\sigma^2}\right) \otimes \mathbf{1} \right], \end{aligned} \quad (24)$$

where  $\mathcal{P} = \mathbf{V} \Lambda$  is a  $N \times n_{ctrl}$  matrix, and the newly kernel weighting  $\widehat{\alpha} = [\alpha_1, \dots, \alpha_{n_{ctrl}}]^T$  is a  $n_{ctrl} \times d$  matrix.

Thus, our proposed spatially constrained Gaussian fields (namely SCGF) can be outlined in Algorithm 1.

## IV. ALGORITHM ANALYSIS AND APPLICATIONS

### A. Algorithm Analysis

In the optimization, we can use a rigid to non-rigid strategy by applying the deterministic annealing technique on the scale parameter  $\sigma$  to improve the algorithm escape the local minimum. More specially, given a large initial value of  $\sigma$  for global rigid transformation, and reducing its value with a fixed annealing rate  $\rho$  towards for local non-rigid transformation by equation  $\sigma = \sigma \times \rho$  iteratively.  $\sigma$  denotes a capture range for each mixture of Gaussian model.  $\beta$  controls the structural strength of the moving point set.  $\beta$  produces locally smooth transformation with small values, while it produces globally translation transformation with large values. In order to get

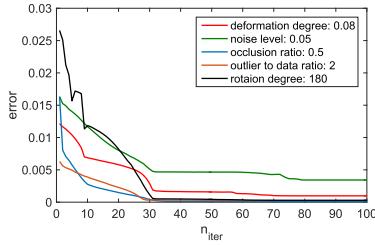


Fig. 1. Convergence experiment on synthesized dataset under the largest degree of deformation, noise, occlusion, outlier, and rotation.

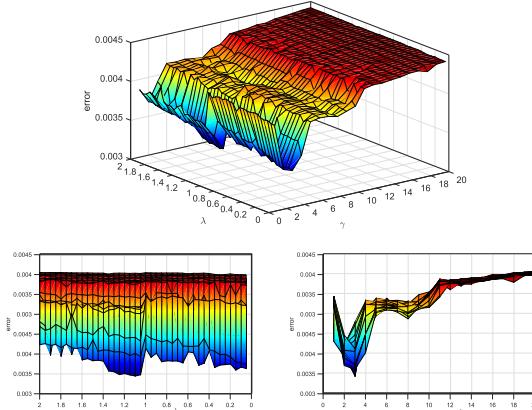


Fig. 2. Model selection of the regularization parameters  $\lambda$  and  $\gamma$  for point set registration.

robust results, the annealing process is need to slow enough for the method. So we set  $\sigma = \det(X^T X/N)^{\frac{1}{2d}}$ ,  $\beta = 0.8$  and  $\rho = 0.98$  throughout this paper. Due to the optimization needing a termination condition, we use the function *fminunc* in Matlab with the options: *{MaxIter=200}*, where the quasi-Newton method uses a mixed quadratic and cubic line search procedure and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula for updating the approximation of the Hessian matrix. In order to reach a good stable local minima, we optimize the correspondences  $C$  and non-rigid transformation  $T$  iteratively, then the experiments show that the SCGF will catch a good stable solution after about 30 iterations, as showed in Fig. 1.

The parameters of regularization terms include  $\lambda$ ,  $\gamma$  which are used to trade-off the smoothness, and the analysis of model selection is shown in Fig. 2. The test result shows that the proposed SCGF performs best for  $\lambda \in [0.1, 1.4]$  and  $\gamma \in [2, 5]$ . In this paper, we fixed them as  $\lambda_1 = 0.1$ ,  $\gamma = 3$ . Note that the construction of the graph Laplacian needs a weight matrix  $W$ , we hereby choose heat kernel with band  $\sigma_s = 0.8$  to define it, and the number of the nearest neighbor is set as  $N_n = 3$ . The Matlab implementation of the graph Laplacian regularization can be available at [http://manifold.cs.uchicago.edu/manifold\\_regularization/manifold.html](http://manifold.cs.uchicago.edu/manifold_regularization/manifold.html) [25].

Typically, the number of control points in  $J$  mainly affects the computational cost of the proposed SCGF for non-rigid point set registration with RBF kernel. Thus, how to choose the number and the positions of control points is intractable. Here, we try a random selection and low-rank approximation methods to determine control points. More precisely,

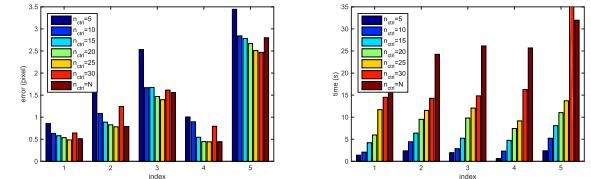


Fig. 3. Registration results on IMM face landmarks under different values of  $n_{ctrl}$  for low-rank kernel matrix approximation. Statistics of registration errors (left) and runtime (right).

## Algorithm 2 Non-Rigid Point Set Registration Using SCGF

```

Input:  $X$ ,  $Y$ .
Output:  $\widehat{X}$ ,  $\alpha^*$ .
Initialize:  $\sigma$ ,  $\beta$ ,  $\lambda$ ,  $\gamma$ ,  $\rho$ ,  $n_{ctrl}$ ,  $k = 0$ ,  $n_{iter} = 30$ ,  $\alpha = 0$ ;
Begin
while  $k < n_{iter}$  do
    Get descriptor  $S_X$  and  $S_Y$  by Eq. (4);
    Assign the correspondences  $C$  by Eq. (5, 6, 7);
    Construct kernel-valued matrix  $\mathcal{K}$  by RBF;
    Compute the transformation parameter  $\alpha$  using SCGF;
    Update the model point set:  $\mathbf{X}' = \mathbf{X} + \mathcal{K}\alpha$ ;
     $\sigma \leftarrow \sigma \times \rho$  (annealing);
     $k = k + 1$ ;
end
Determine  $\widehat{X}$  by the optimal parameter  $\alpha^*$ ;
return  $\widehat{X}$  and  $\alpha^*$ .

```

low-rank kernel-valued matrix approximation is applied to reduce the computational complexity, the important parameter  $n_{ctrl}$  denotes the number of the selected eigenvalues, and as a trade-off, it can be tuned according to the registration accuracy and computing cost. The balance experiment shows that the proposed SCGF performs better for  $n_{ctrl} = [15, 20]$  (where  $N = 105$ ,  $d = 2$ ), and as shown in Fig. 3.

The computational complexity of correspondences weight matrix assignment is  $\mathcal{O}(N^3)$ , and computing non-rigid transformation optimization needs almost  $\mathcal{O}(N^3 + N^2 + MN)$ . Precisely, the Hungarian method is used to match point set  $X$  and  $Y$  in  $\mathcal{O}(N^3)$  runtime, while the optimization needs  $\mathcal{O}(N^2 + MN)$  for computing the objective function and its gradient. Moreover, by using the low-rank approximation strategy, if  $n_{ctrl} \ll N$ , the computational complexity reduces from  $\mathcal{O}(N^3 + N^2 + MN)$  down to  $\mathcal{O}(n_{ctrl}^3 + n_{ctrl}^2 + Mn_{ctrl})$ , then the computational complexity of the proposed SCGF is approximated as  $\mathcal{O}(MN)$ .

Note that 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.

## B. Applications

Non-rigid point set registration using the proposed SCGF is outlined in Algorithm 2.

Moreover, the proposed SCGF can be used to estimate the transformation robustly from the putative correspondences

**Algorithm 3** Correspondences Rejection Using SCGF

---

**Input:**  $C_0, X, Y$ .  
**Output:**  $\hat{C}$ .  
**Initialize:**  $\sigma, \beta, \lambda, \gamma, \rho, n_{ctrl}, \alpha = 0$ ;  
**Begin**  
    Construct kernel-valued matrix  $\mathcal{K}$  by RBF;  
    Compute the transformation parameter  $\alpha$  using SCGF;  
     $\widehat{X} = X + \mathcal{K}\alpha$ ;  
    Compute the index  $\mathcal{I}$  by Eq. (25);  
    Determine  $\hat{C} = C_0(\mathcal{I})$ ;  
    **return**  $\hat{C}$ ;  
**End**

---

which contains incorrect matches, then the SCGF can be applied for correspondences rejection (Part of our previous work for correspondences rejection under manifold regularization has been reported in [27]). More precisely, the correct correspondences are identified by comparing the transformed model  $\widehat{X}$  point set and the scene  $Y$  with a threshold  $\eta$ . Then the index  $\mathcal{I}$  of the recovered matches can be defined as:

$$\mathcal{I} = \exp\left(-\frac{\text{diag}[(Y - \widehat{X})(Y - \widehat{X})^T]}{\sigma^2}\right) \geq \eta, \quad (25)$$

where  $\eta \in (0, 1)$  is a rejecting threshold which controls the balance between precision and recall values. Briefly, correspondences rejection using SCGF is outlined in Algorithm 3.

## V. EXPERIMENTS

The SCGF is implemented in Matlab and tested on an Intel Core i5 CPU 2.5GHz with 8GB RAM. The source code of our proposed method can be available from <https://sites.google.com/site/2013gwang/SCGF.zip>.

In our experiments, we tested two scenarios: 1) non-rigid point set registration for two-dimension and three-dimension data, and 2) correspondences rejection.

For a comprehensive comparison, the proposed SCGF is compared with some state-of-the-art methods, and extensive experimental results are presented to demonstrate the superiority of the proposed SCGF on the task of point set registration. Precisely, six point set registration approaches, and four correspondences rejecting approaches are included in our comparative study: TPS-RPM [3], SC [23], IDSC [21], QPCCP [16], GMMReg [6], CPD [8], and RPM-L2E [12]; Random Sample Consensus (RANSAC) [28], Identify Correspondence Function (ICF) [29] based on the support vector regression, non-rigid RANSAC [30], and Vector Field Consensus (VFC) [31]. All methods are implemented in Matlab and tested in the same environment.

### A. Results on Synthesized Point Set

This dataset is constructed by Chui and Rangarajan [3], 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,

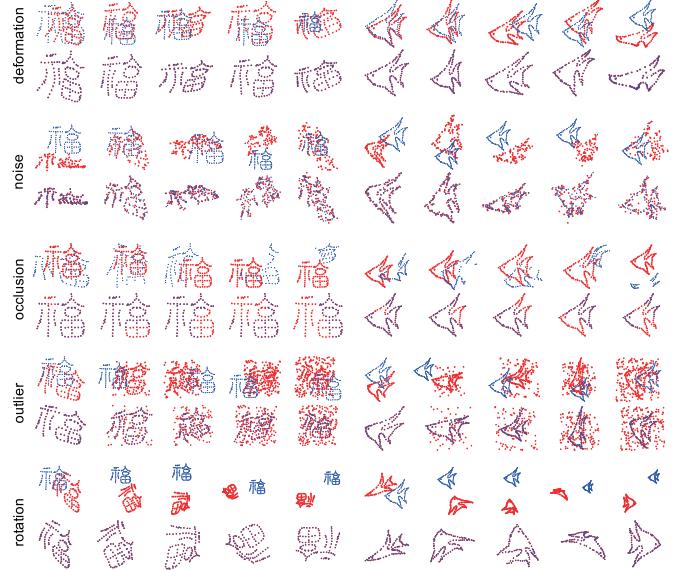


Fig. 4. Registration results of the proposed SCGF algorithm on the synthesized data: Chinese character and fish shape. From top to bottom: deformation, noise, occlusion, outliers, and rotation. From left to right in each group, the gradation level becomes larger.

i.e., deformation, noise, occlusion, outliers, and rotation are used to evaluate the accuracy and robustness of registration approaches, resulting in 5000 pairs of point set in total.

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 registration 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 ('+') are all well aligned onto the scene sets ('o') except the scene sets are distorted 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 error of several non-rigid registration methods using the root-mean-square error (RMSE) on the synthesized data, where  $\text{RMSE} = \sqrt{\sum(\|Y - \widehat{X}\|^2)/N}$ . Quantitative experimental comparison results demonstrate that the proposed SCGF 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 SCGF with global to local regularization refinement helps to estimate non-rigid transformations robustly.

Table I 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  $N = 100, M = 100$  points.

In order to state the contribution of the spatially constraint (i.e., graph Laplacian regularization), we set up a comparative experiment by adding the graph Laplacian regularization term on SC [23] and IDSC [21] respectively. Fig. 6 shows the comparison results by computing the accuracy and recall,

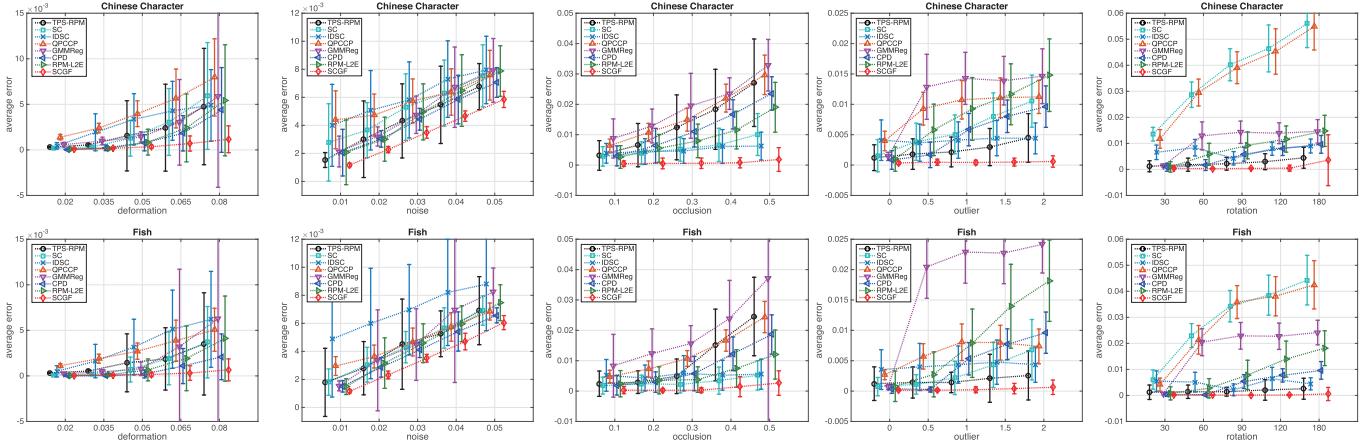


Fig. 5. Comparison between six methods and our SCGF to register point sets on Chinese character and fish shape. (a) Registration errors plot on Chinese character data. (b) Registration errors plot on fish shape.

TABLE I

THE AVERAGE RUNTIME (SEC.) OF THE REGISTRATION ON SYNTHESIZED DATA (*Chinese Character*) WITH 80 ITERATIONS. FROM TOP ROW TO BOTTOM, THE DEGREE LEVEL BECOMES LARGER FROM 1 TO 5

	deformation	noise	occlusion	outlier	rotation
1	15.05	16.70	15.37	16.01	15.99
2	16.69	19.14	15.81	20.08	16.80
3	16.95	19.99	15.44	23.7	17.07
4	17.44	20.17	15.30	27.25	16.54
5	17.80	19.53	15.92	29.00	17.42

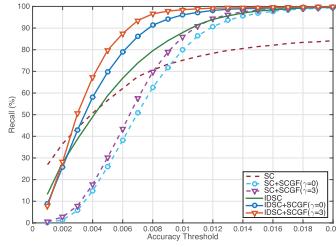


Fig. 6. Comparison between SC and IDSC with SCGF ( $\lambda = 0, 3$ ) on the whole Chinese character and fish shape datasets.

where the accuracy and recall metric curves described in [32]. Though we can see that IDSC outperforms SC, our proposed SCGF with  $\lambda = 3$  can improve the performance of SC and IDSC without adding the graph Laplacian regularization term (i.e., SCGF with  $\lambda = 0$ ).

### B. Results on IMM Face Landmarks

The IMM database (<http://www.imm.dtu.dk/~aam/datasets/datasets.html>) consists of face landmarks with expression and multi-view changes. 58 point landmarks are sampled from a face with different facial expressions and poses.

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. 7. The non-rigid deformation of the scene set becomes large with increasing degree of viewpoint and changing facial expression. The qualitative registration results by aligning the

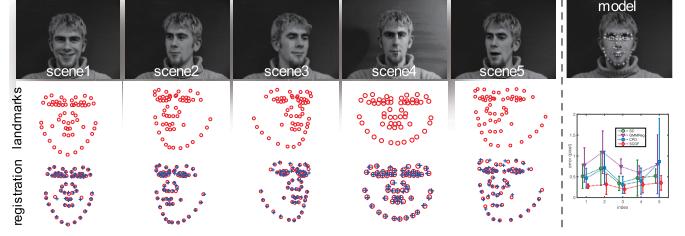


Fig. 7. Registration results on face landmarks. From top to bottom: the model landmarks, the scene landmarks, and the registration results of the proposed method. The rightmost figure is the comparison analysis, and the error bars indicate the registration error means and standard deviations over five samples in each group of data.

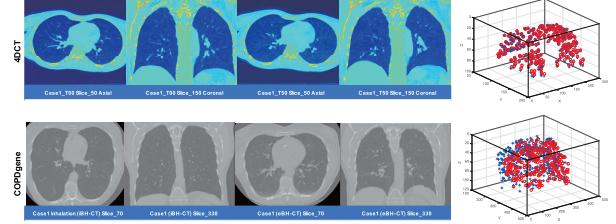


Fig. 8. The example of 4DCT and COPDgene data. The rightmost plots denote the landmarks which are labeled by experts for the corresponding lung data (left).

model set onto the five scene sets respectively as shown. We can see that the point sets are well aligned by the SCGF algorithm, however, the result on scene 5 is slightly bad due to the sampled landmarks are contaminated by noise when experiencing large posture and expression changes. As shown in the rightmost figure of Fig. 7, the proposed SCGF can get the better registration performance than the well-known SC [23], GMMReg [6], and CPD [8] methods on five groups of faces, where the shape context registration method uses the TPS transformation model.

### C. Results on 3D Point Set

As point set registration plays a key role in medical image registration [33], [34]. In this experiment, we use the DIR-lab (<http://www.dir-lab.com/index.html>) data which

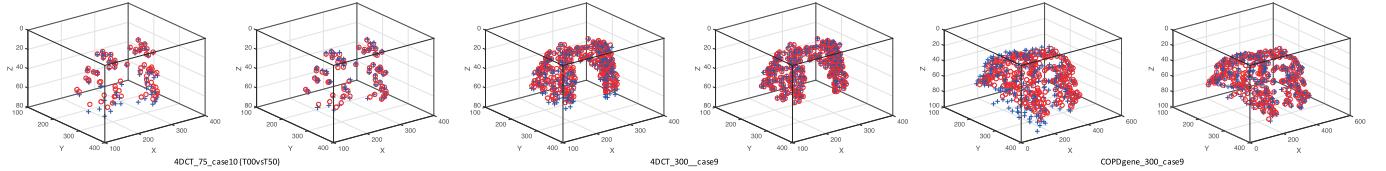


Fig. 9. The example of the registration results of the 4DCT and COPDgene data using SCGF. In each group, the left plot denotes the initialization, and the right one denotes the registration.

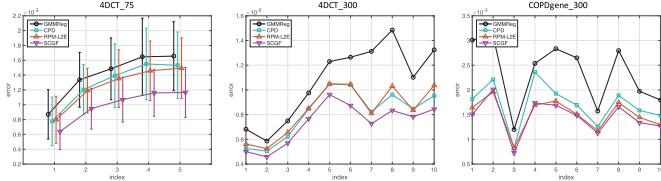


Fig. 10. The comparison results using average error (RMSE). The index of the first “4DCT\_75” denotes the average error on {“T00vsT10”, “T00vsT20”, “T00vsT30”, “T00vsT40”, “T00vsT50”} in the whole cases (case1 to case10), while each index of the “4DCT\_300”, and “COPDgene\_300” denote the corresponding case.

provides currently two sets of benchmark data [35]. Precisely, the first 4DCT contains ten (“case1” to “case10”) thoracic 4D CT images which consist of ten 3D CT scans, which are used to treat thoracic tumors. The second COPD dataset originates from the National Heart Lung Blood Institute COPDgene study archive and contains ten inspiratory/expiratory breath-hold 3D CT image pairs [36], [37]. Each data set contains the publicly available anatomical landmark pairs which are labeled by medical experts manually and are often located at prominent bifurcations of the bronchial tubes. We use the annotated landmarks to evaluate the registration approaches quantitatively. The sample images and the landmarks are shown in Fig. 8, where the axial and coronal views are plotted.

Based on the available landmarks, we design three experiments to implement the comparison study. Especially, we set “caseID\_300\_T00\_xyz” as the model point set, and “caseID\_300\_T50\_xyz” is set as the scene, then we can obtain ten pairs of point sets, where “ID” $\in$ [1, 10]. In addition, we set “4DCT-75\_T00” as the model and the other four (“T10” to “T50”) are the scene point sets in each case, then we get 50 pairs of point sets. For COPD dataset, we are assuming that the inspiratory case is set as the model, while the expiratory case for the scene. In all, 70 pairs of 3D lung landmark point sets are constructed for the experiments. Note that these three test datasets are named “4DCT\_75”, “4DCT\_300”, and “COPDgene\_300”, respectively.

In Fig. 9, we show three samples of the registration results (without normalization) using the proposed SCGF, where the model point sets are aligned onto the corresponding scene point sets. More specially, the inspiratory/expiratory breath-hold cases in COPDgene have a higher degree of deformation, and it is relatively intractable to deal with.

After the above qualitative evaluation, we also show the statistic of the comparison results (with normalization) using registration error (RMSE) in Fig. 10. As depicted in this figure, our proposed method achieves the best objective performance

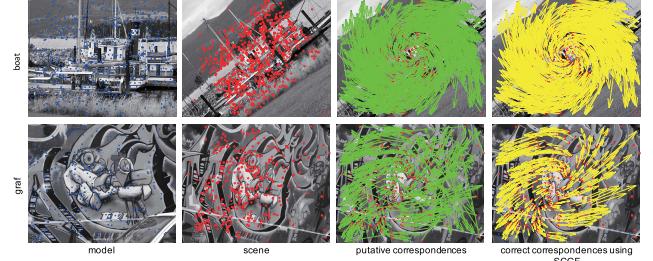


Fig. 11. The example of the correspondences rejection results using SCGF. The feature points are shown in each image, and the green arrows denote the putative matches, the yellow ones in the rightmost plot denote the correct correspondences after removing outliers via SCGF.

among the compared approaches. More precisely, the SCGF method gives the lowest registration error for all tested cases. The error bars of “4DCT\_75” denote the average registration error and standard deviation over ten cases. In all, “4DCT\_75” consists of 50 pairs of 4DCT landmark sets. As the complexity of the inspiratory/expiratory COPDgene, the errors on this data are slightly higher than 4DCT data.

#### D. Results on Correspondences Rejection

Correspondence recovering is a hard problem in point set registration. Naturally, not all estimated correspondences by the feature matching are correct in most experiments. Since incorrect correspondences can negatively affect the final estimation of the non-rigid transformation, they need to be rejected [38].

The Oxford image dataset is used to test the correspondences rejection performance of the proposed SCGF with the putative correspondences parameter  $C_0$ . These affine covariant region datasets include bark (zoom & rotation), bikes (blur), boat (zoom & rotation), graf (multi-view), leuven (light change), trees (blur), ubc (compression), and wall (multi-view). The data sets are available at <http://www.robots.ox.ac.uk/vgg/data/data-aff.html>. For the quantitative evaluation, we use the accuracy  $acc = (tp + tn)/(tp + tn + fp + fn)$ , precision  $pre = tp/(tp + fp)$ , and recall  $rec = tp/(tp + fn)$  as the evaluation metrics, where the terms  $tp$ ,  $tn$ ,  $fp$ ,  $fn$  denote true positives, true negatives, false positives, and false negatives, respectively. We compare SCGF with five robust point matching approaches: RANSAC, ICF, CPD, VFC, and non-rigid RANSAC. We use the VLFeat [39] toolbox to extract the feature points of each image, and the putative matches are generated by the nearest neighbor matching. Fig. 11 shows the example of the results using SCGF, where the feature points are labeled in the model and scene images, and the putative correspondences

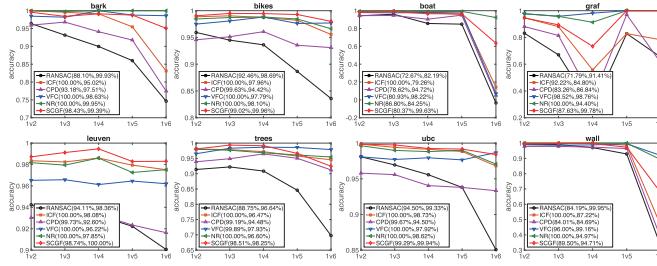


Fig. 12. Comparison between correspondences rejection methods and our SCGF to remove mismatches for robust point matching, where the results using precision and recall pairs (*pre*, *rec*) are also showed in the legend.

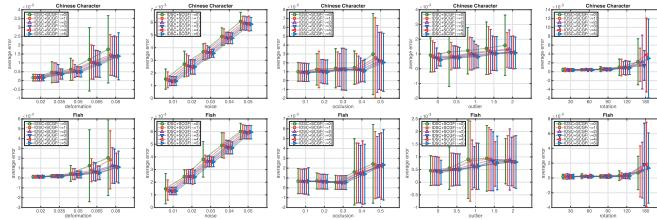


Fig. 13. The comparison study for the regularization parameter  $\gamma$  selection in SCGF. The X-axis denotes the degradation which becomes larger from left to right.

are obtained from the initial matching. The quantitative evaluations are shown in Fig. 12. From image pair “1v2” to “1v6” in Fig. 12, the correspondence rejection becomes hard to solve.

#### E. Discussion for Model Selection

In this experiment, we discuss for the Laplacian regularization parameter  $\gamma$  briefly. Due to the normalization processing makes the registration error very small in the synthesized datasets, we can see that the results using SCGF with  $\gamma = \{1, 2, 3, 4, 5\}$  are better than those with  $\gamma = 0$ , and the standard deviations of SCGF are less than the results by SCGF with  $\gamma = 0$  in most scenarios, as shown in Fig. 13. More specially, the graph Laplacian regularization term plays an important role when facing a large degree of degradations.

## VI. CONCLUSIONS

In this paper, we present a robust non-rigid transformation estimation algorithm, called spatially constrained Gaussian fields (SCGF), and its applications for robust non-rigid point set registration and correspondences rejection. We utilize the graph Laplacian regularization to constrain the geometry structure of the data and address the ill-posed problem in transformation estimation. The proposed method uses an estimator based on Gaussian fields instead of building a more complex model that includes inliers and outliers to get more robust performance. Moreover, the proposed SCGF for non-rigid point set registration is subject to global rigid and local non-rigid geometric constraints, where a graph laplacian regularized term is added to preserve the intrinsic geometry of the transformed set. SCGF uses a robust objective function and the quasi-Newton algorithm to estimate the likely correspondences, and the non-rigid transformation parameters iteratively. Meanwhile, the deterministic annealing

method provides a coarse-to-fine scaling strategy which is used to propagate local structure and escape local minimums. Experimental results demonstrate that much more accuracy can be achieved using SCGF than the state-of-the-art methods on both two-dimension and three-dimension data.

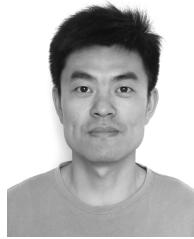
## ACKNOWLEDGMENTS

The authors also would like to thank the anonymous reviewers for their constructive suggestions which helped them improve the manuscript.

## REFERENCES

- [1] B. Schölkopf, R. Herbrich, and A. J. Smola, “A generalized representer theorem,” in *Proc. Int. Conf. Comput. Learn. Theory*, 2001, pp. 416–426.
- [2] P. J. Besl and D. N. McKay, “A method for registration of 3-D shapes,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.
- [3] H. Chui and A. Rangarajan, “A new point matching algorithm for non-rigid registration,” *Comput. Vis. Image Understand.*, vol. 89, nos. 2–3, pp. 114–141, Feb. 2003.
- [4] Y. Zheng and D. Doermann, “Robust point matching for nonrigid shapes by preserving local neighborhood structures,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 643–649, Apr. 2006.
- [5] Y. Tsin and T. Kanade, “A correlation-based approach to robust point set registration,” in *Proc. Comput. Vis.-(ECCV)*, 2004, pp. 558–569.
- [6] B. Jian and B. C. Vemuri, “Robust point set registration using Gaussian mixture models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1633–1645, Aug. 2011.
- [7] A. L. Yuille and N. M. Grzywacz, “The motion coherence theory,” in *Proc. IEEE 2nd Int. Conf. Comput. Vis. (ICCV)*, Dec. 1988, pp. 344–353.
- [8] A. Myronenko and X. Song, “Point set registration: Coherent point drift,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.
- [9] L. Greengard and J. Strain, “The fast gauss transform,” *SIAM J. Sci. Statist. Comput.*, vol. 12, no. 1, pp. 79–94, 1991.
- [10] J. Ma, J. Zhao, and A. L. Yuille, “Non-rigid point set registration by preserving global and local structures,” *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 53–64, Jan. 2016.
- [11] H. Li, T. Shen, and X. Huang, “Global optimization for alignment of generalized shapes,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 856–863.
- [12] J. Ma, J. Zhao, J. Tian, Z. Tu, and A. L. Yuille, “Robust estimation of nonrigid transformation for point set registration,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2013, pp. 2147–2154.
- [13] J. Ma, W. Qiu, J. Zhao, Y. Ma, A. L. Yuille, and Z. Tu, “Robust  $L_2 E$  estimation of transformation for non-rigid registration,” *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1115–1129, Mar. 2015.
- [14] G. Wang, Z. Wang, W. Zhao, and Q. Zhou, “Robust point matching using mixture of asymmetric Gaussians for nonrigid transformation,” in *Computer Vision—ACCV*. Cham, Switzerland: Springer, 2015, pp. 433–444.
- [15] G. Wang, Z. Wang, Y. Chen, and W. Zhao, “A robust non-rigid point set registration method based on asymmetric Gaussian representation,” *Comput. Vis. Image Understand.*, vol. 141, pp. 67–80, Dec. 2015.
- [16] W. Lian, L. Zhang, Y. Liang, and Q. Pan, “A quadratic programming based cluster correspondence projection algorithm for fast point matching,” *Comput. Vis. Image Understand.*, vol. 114, no. 3, pp. 322–333, 2010.
- [17] F. Boughorbel, M. Mercimek, A. Koschan, and M. Abidi, “A new method for the registration of three-dimensional point-sets: The Gaussian fields framework,” *Image Vis. Comput.*, vol. 28, no. 1, pp. 124–137, 2010.
- [18] J. Ma, J. Zhao, Y. Ma, and J. Tian, “Non-rigid visible and infrared face registration via regularized Gaussian fields criterion,” *Pattern Recognit.*, vol. 48, no. 3, pp. 772–784, 2015.
- [19] G. Wang, Z. Wang, Y. Chen, Q. Zhou, and W. Zhao, “Context-aware Gaussian fields for non-rigid point set registration,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5811–5819.
- [20] D. A. Murio, *The Mollification Method and the Numerical Solution of Ill-Posed Problems*. Hoboken, NJ, USA: Wiley, 2011.

- [21] H. Ling and D. W. Jacobs, "Shape classification using the inner-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.
- [22] A. Frome, D. Huber, R. Kolluri, T. Bülow, and J. Malik, "Recognizing objects in range data using regional point descriptors," *Lect. Notes Comput. Sci.*, vol. 3023, pp. 224–237, 2004.
- [23] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.
- [24] L. Rabiner, "Combinatorial optimization: Algorithms and complexity," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 32, no. 6, pp. 1258–1259, Dec. 1984.
- [25] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," *J. Mach. Learn. Res.*, vol. 7, pp. 2399–2434, Nov. 2006.
- [26] F. R. Chung, *Spectral Graph Theory*. Providence, RI, USA: AMS, 1997.
- [27] G. Wang, Z. Wang, Y. Chen, X. Liu, Y. Ren, and L. Peng, "Learning coherent vector fields for robust point matching under manifold regularization," *Neurocomputing*, vol. 216, pp. 393–401, Dec. 2016.
- [28] M. A. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [29] X. Li and Z. Hu, "Rejecting mismatches by correspondence function," *Int. J. Comput. Vis.*, vol. 89, no. 1, pp. 1–17, 2010.
- [30] 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 *Comput. Vision–ECCV*. Springer, 2012, pp. 274–287.
- [31] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014.
- [32] J. Starck and A. Hilton, "Correspondence labelling for wide-timeframe free-form surface matching," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
- [33] D. Shen and C. Davatzikos, "Hammer: Hierarchical attribute matching mechanism for elastic registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 11, pp. 1439–1421, Nov. 2002.
- [34] G. Wang, Z. Wang, Y. Chen, and W. Zhao, "Robust point matching method for multimodal retinal image registration," *Biomed. Signal Process. Control*, vol. 19, pp. 68–76, Apr. 2015.
- [35] R. Castillo *et al.*, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," *Phys. Med. Biol.*, vol. 54, no. 7, pp. 1849–1870, 2009.
- [36] R. Castillo *et al.*, "A reference dataset for deformable image registration spatial accuracy evaluation using the copdgene study archive," *Phys. Med. Biol.*, vol. 58, no. 9, pp. 2861–2877, 2013.
- [37] S. Hermann, "Evaluation of scan-line optimization for 3D medical image registration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2014, pp. 205–209.
- [38] G. Wang, Z. Wang, Y. Chen, Q. Zhou, and W. Zhao, "Removing mismatches for retinal image registration via multi-attribute-driven regularized mixture model," *Inf. Sci.*, vol. 372, pp. 492–504, Dec. 2016.
- [39] A. Vedaldi and B. Fulkerson, "VLfeat: An open and portable library of computer vision algorithms," in *Proc. 18th ACM Int. Conf. Multimedia*, 2010, pp. 1469–1472.



**Gang Wang** received the Ph.D. degree in computer science and technology from Tongji University in 2016. He is currently an Assistant Professor with the School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China. His current research interests include computer vision, pattern recognition, and data analysis.



**Qiangqiang Zhou** received the B.S. degree from Jiangxi Normal University, Nanchang, China, in 2003, and the M.S. degree in Communication and Information System from Central South University, Changsha, China, in 2009. He is currently pursuing the Ph.D. degree with the CAD Research Center of Tongji University, Shanghai, China. His research interests include pattern recognition, machine learning, and computer vision.



**Yufei Chen** received the Ph.D. degree from Tongji University in 2010. She is currently a Senior Lecturer with the CAD Research Center of Tongji University, Shanghai, China. She was a Post-Doctoral Researcher with the Control Science and Engineering of Tongji University from 2010 to 2012. She was also a Guest Researcher with the Fraunhofer Institute for Computer Graphics Research, Germany, from 2008 to 2009. Her research topics include image processing and data analysis.