

Non-rigid Point Set Registration with Global-Local Topology Preservation

Song Ge, Guoliang Fan* and Meng Ding

School of Electrical and Computer Engineering
Oklahoma State University, Stillwater, OK 74078

{song.ge, guoliang.fan, meng.ding}@okstate.edu

Abstract

We propose a new point set registration method, *Global-Local Topology Preservation* (GLTP), which can cope with complex non-rigid transformations including highly articulated deformation. The registration is formulated as a Maximum Likelihood (ML) estimation problem with two topologically complementary constraints. The first is the previous Coherent Point Drift (CPD) that encodes a global topology constraint by moving one point set coherently to align with the second set. The second, which is inspired by the idea of Local Linear Embedding (LLE), is introduced to handle highly articulated non-rigid deformation while sustaining the local structure. Without any pre-segmentation, the newly introduced LLE constraint is particularly useful and effective when there are multiple non-coherent and non-rigid local deformations (i.e., the CPD assumption may be violated). We have derived the EM algorithm for the ML optimization constrained with both CPD and LLE terms, leading to the new GLTP algorithm. Experimental results on 2D and 3D examples show its accuracy and robustness in the presence of outliers and noise, especially in the case of highly-articulated non-rigid transformation.

1. Introduction

Point set registration is a fundamental issue for many computer vision applications. Particularly, due to the emergence of more affordable depth sensors, it is becoming a more prevalent and intriguing research topic, e.g., human pose estimation [21] or body shape modeling [20]. The registration techniques usually fall into two categories: rigid and non-rigid depending on the the underlying transformation model. Non-rigid registration which is more common and desirable in practice remains a challenge in computer vision. Especially, the need for general point set registration that supports non-rigid and articulated transformations is still under-served in current literature.

*This work is supported by Oklahoma Center for the Advancement of Science and Technology (OCAST) under grants HR09-030 and HR12-30.

Iterative Closet Point (ICP) [3, 22] is a classic rigid registration method which iteratively assigns correspondence and then finds the least square transformation by using the estimated correspondence. The shape context descriptor proposed in [2] was used in [23, 13] to initialize correspondence for non-rigid registration. A graph was developed in [23] to represent the local neighborhood relationship of each point, and then registration was converted to a graph matching problem. A robust method was proposed in [13] to estimate the non-rigid transformation defined in a reproducing kernel Hilbert space (RKHS). Recently, Gaussian Mixture Model (GMM) based probabilistic registration algorithms are becoming an important category, where the main idea is to represent the point sets by a density function and then registration is cast as a density estimation problem under different transformation models. For example, thin-plate splines (TPS) was used to model the non-rigid transformation in [5, 6]. A Gaussian radial basis functions (GRBF) based displacement function with a global smoothness constrain, Coherent Point Drift (CPD), was introduced in [16, 15]. The articulated rigid transformation was used in [7] where the point set is pre-segmented based on a kinematic chain. In [11, 10], GMM-based registration is treated as an alignment between two distributions corresponding to two point sets.

We propose a new GMM-based point set registration algorithm, called Global-Local Topology Preservation (GLTP), which can deal with general non-rigid transformation including highly articulated deformation, such as 3D human data. Existing algorithms involve locally rigid assumption which normally requires pre-segmentation [18, 7] or some initial conditions, such as known correspondence [12, 14, 13] and nonsignificant articulated deformation [21], to cope with this challenge. GLTP is more flexible and general which incorporates two topologically complementary constraints, CPD-based global coherence and LLE-based local constraint, in a unified maximum likelihood (ML) estimation framework. A balanced treatment between CPD and LLE in the EM-based ML optimization allows GLTP to handle complex non-rigid articulated transformations, which cannot be handled by most existing approaches.

2. GMM-based Point Set Registration

Consider the point sets $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M]^T$, $\mathbf{y}_m \in R^D$ as the template with a sparse distribution and $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N]^T$, $\mathbf{x}_n \in R^D$ as the target data with a dense distribution. We denote the spatial transformation as $\mathcal{T}(\mathbf{Y}, \Theta)$, which can be considered as a function of \mathbf{Y} with parameters Θ . To account for outliers in \mathbf{X} , a uniform component with weight ω , $0 \leq \omega \leq 1$, is added [15, 5]. Assuming all the Gaussian components are independently distributed with the same isotropic variance σ^2 and weight, the joint GMM probability density function can be written as

$$p(\mathbf{X}) = \prod_{n=1}^N p(\mathbf{x}_n) = \prod_{n=1}^N \sum_{m=1}^{M+1} \pi_m p(\mathbf{x}_n | m), \quad (1)$$

where $p(\mathbf{x}|m) = \frac{1}{(2\pi\sigma^2)^{\frac{D}{2}}} \exp(-\frac{\|\mathbf{x} - \mathcal{T}(\mathbf{y}_m, \Theta)\|^2}{2\sigma^2})$ ($m = 1, \dots, M$), $p(\mathbf{x}|m) = \frac{1}{N}$ for $m = M + 1$, $\pi_m = \frac{1-\omega}{M}$ ($m = 1, \dots, M$), and $\pi_{M+1} = \omega$. Then registration is converted to the problem of finding Θ and σ^2 that minimizes the negative log-likelihood of (1). Following the EM algorithm for GMM-based clustering [4, 15], we can find the objective (E-step) as:

$$Q(\Theta, \sigma^2) = \sum_{n=1}^N \sum_{m=1}^M p^{old}(m|\mathbf{x}_n) \frac{\|\mathbf{x}_n - \mathcal{T}(\mathbf{y}_m, \Theta)\|^2}{2\sigma^2} + \frac{N_p D}{2} \ln(\sigma^2), \quad (2)$$

where $N_p = \sum_{n=1}^N \sum_{m=1}^M p^{old}(m|\mathbf{x}_n)$ and $p^{old}(m|\mathbf{x}_n)$ are the posterior probabilities that can be computed using the old GMM parameters based on the Bayes rule as:

$$p^{old}(m|\mathbf{x}_n) = \frac{\exp(-\frac{1}{2} \|\frac{\mathbf{x}_n - \mathcal{T}(\mathbf{y}_m, \Theta)}{\sigma^{old}}\|^2)}{\sum_{i=1}^M \exp(-\frac{1}{2} \|\frac{\mathbf{x}_n - \mathcal{T}(\mathbf{y}_i, \Theta)}{\sigma^{old}}\|^2) + \frac{(2\pi\sigma^2)^{\frac{D}{2}} \omega M}{(1-\omega)N}}. \quad (3)$$

Then in the M-step we can find the new GMM parameters by minimizing (2). The EM algorithm performs iteratively by alternating between E-step and M-step until it converges.

In order to ensure accurate correspondence estimation during non-rigid registration, appropriate regularization is essentially needed. So we can extend (2) by adding a regularization term to get a constrained optimization problem:

$$Q_{\mathcal{R}}(\Theta, \sigma^2) = Q(\Theta, \sigma^2) + \mathcal{R}(\mathcal{T}(\mathbf{y}_m, \Theta)), \quad (4)$$

where $\mathcal{R}(\mathcal{T}(\mathbf{y}_m, \Theta))$ denotes the regularization on the non-rigid transformation. Therefore, different GMM-based algorithms could be discussed collectively based on the type of non-rigid transformations and the choice of the regularization term.

3. Proposed Work

3.1. Motivation

In this work, we study GMM-based non-rigid point set registration, especially, *non-rigid articulated registration*, where transformations are non-rigid both globally and locally, unlike those dealing with rigid articulated registration (e.g., [7]). This case is more general and flexible to accommodate multi-part non-rigid deformations along with the global articulation observed in many real-world cases, such as hand tracking and human pose estimation, where parts are often assumed to be locally rigid for simplicity. Existing non-rigid registration algorithms, such as CPD or TPS are not effective to handle non-rigid articulated transformations. Fig. 1 column (a) shows a template human model (first row) and a target 3D human data (second row). The two point sets have different body shapes (non-rigid local deformations) and poses (global articulation). The template model was manually color-labeled to indicate the estimated correspondence after registration. It is worth mentioning that those point labels are not used during registration and they are only for visualization purpose.

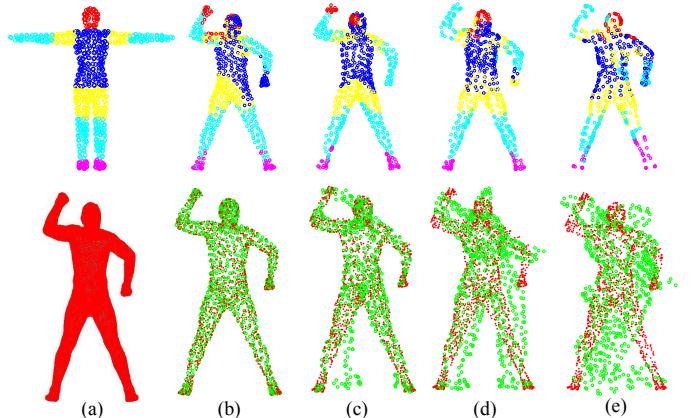


Figure 1. CPD registration results under different β values: column (a) the template where different body parts are labeled by different colors (first row) and the target data (second row); Registration results (b)-(e) with respect to $\beta = 0.1, 0.5, 2, 10$ where the first and second rows show the estimated correspondence and the matching result respectively.

TPS failed in this 3D example, and CPD performs also poorly, as shown in Fig. 1 column (b) to (e) (the first row shows correspondence estimation and second row shows registration results) respect to different β values which are the width of the Gaussian filter and control local regularization. The main reason that CPD was successfully applied to human pose estimation in [21] is due to a good pose initialization by finding the best match in a large training data set that reduces the articulation deformation between two point sets to be registered.

3.2. Global-Local Topology Constraints

Our goal is to achieve robust and accurate non-rigid highly-articulated registration. The key is to preserve both the global connectivity and the local structure. We want to incorporate two topologically complementary terms in a unified optimization process. For the global topology constraint, we still use the idea of CPD where the non-rigid transformation is defined as a GRBF-based displacement function derived by using calculus of variation:

$$\mathcal{T}(\mathbf{Y}, \mathbf{W}) = \mathbf{Y} + \mathbf{G}\mathbf{W}, \quad (5)$$

where $\mathbf{G}_{M \times M}$ is the Gaussian kernel matrix with element $g_{ij} = \exp(-\frac{1}{2} \|\frac{\mathbf{y}_i - \mathbf{y}_j}{\beta}\|^2)$ and $\mathbf{W}_{M \times D}$ is the weight matrix of the Gaussian kernel. Then we can regularize the weight matrix \mathbf{W} to enforce the motion coherence and consequentially to preserve the global topology:

$$E_{global}(\mathbf{W}) = \text{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{W}). \quad (6)$$

The global topology is useful to keep the overall spatial connectivity of the point set during registration. However, it could be counter-productive to handle strong articulation and detailed local deformation. We need a local topology term that is aimed to preserve the local neighborhood structure and to accommodate non-rigid articulated deformation. Our work is deeply inspired by [19] where the LLE algorithm was proposed as a nonlinear dimensionality reduction method to preserve the local neighborhood structure in the low-dimensional latent space. In the context of registration, we hope that local structures in \mathbf{Y} are preserved after non-rigid transformation. We have three steps to apply the LLE idea for non-rigid registration. First, compute the K nearest neighbors of each point in \mathbf{Y} according the Euclidean distance. Second, represent each point in \mathbf{Y} by a weighted linear combination of its neighbors. The weights can be obtained by minimizing the cost function

$$E_{LLE}(\mathbf{L}) = \sum_{m=1}^M \|\mathbf{y}_m - \sum_{i=1}^K \mathbf{L}_{mi} \mathbf{y}_i\|^2, \quad (7)$$

where the $M \times K$ matrix \mathbf{L} contains the neighborhood information for each point in \mathbf{Y} . Third, compute the Gaussian kernel weight matrix \mathbf{W} that is used by CPD to represent non-rigid transformation for each point \mathbf{y}_m , so that \mathbf{W} allows each point to be reconstructed by its neighbors with weight \mathbf{L}_{mi} after the non-rigid transformation. \mathbf{W} can be estimated by minimizing the cost function:

$$E_{local}(\mathbf{W}) = \sum_{m=1}^M \|(\mathbf{y}_m + \mathbf{G}(m, \cdot) \mathbf{W}) - \sum_{i=1}^K \mathbf{L}_{mi} (\mathbf{y}_i + \mathbf{G}(m, \cdot) \mathbf{W})\|^2, \quad (8)$$

where $\mathbf{G}(m, \cdot)$ denotes the m^{th} row of \mathbf{G} .

The underlying assumption of LLE-based local topology term is *local deformation coherence* which is complementary to that of CPD, i.e., *global drift coherence*. Using both LLE and CPD terms will allow us to preserve the local and global topologies for complex non-rigid and highly articulated point set registration.

3.3. GLTP-based Registration Algorithm

Our proposed GLTP method treats point set registration problem as a GMM density estimation problem. Based on the general formulation in (2) and given the two regularization terms to preserve the global and local topologies, the objective function of GLTP can be written in the form

$$\begin{aligned} Q_{GLTP}(\mathbf{W}, \sigma^2) = & \sum_{m,n=1}^{M,N} p^{old}(m|\mathbf{x}_n) \frac{\|\mathbf{x}_n - (\mathbf{y}_m + \mathbf{G}(m, \cdot) \mathbf{W})\|^2}{2\sigma^2} \\ & + \frac{N_p D}{2} \ln(\sigma^2) + \frac{\alpha}{2} E_{global}(\mathbf{W}) + \frac{\lambda}{2} E_{local}(\mathbf{W}), \end{aligned} \quad (9)$$

where α and λ are the two trade-off parameters among the GMM matching term and topological constraint terms. Then we rewrite the objective function (9) in a matrix form as:

$$\begin{aligned} Q_{GLTP}(\mathbf{W}, \sigma^2) = & \frac{1}{2\sigma^2} \{ \text{Tr}(\mathbf{X}^T d(\mathbf{P}^T \mathbf{1}) \mathbf{X}) \\ & - 2 \text{Tr}(\mathbf{Y}^T \mathbf{P} \mathbf{X}) - 2 \text{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{P} \mathbf{X}) + \text{Tr}(\mathbf{Y}^T d(\mathbf{P} \mathbf{1}) \mathbf{Y}) \\ & + 2 \text{Tr}(\mathbf{W}^T \mathbf{G} d(\mathbf{P} \mathbf{1}) \mathbf{Y}) + \text{Tr}(\mathbf{W}^T \mathbf{G} d(\mathbf{P} \mathbf{1}) \mathbf{G} \mathbf{W}) \} \\ & + \frac{N_p D}{2} \ln(\sigma^2) + \frac{\alpha}{2} \text{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{W}) + \frac{\lambda}{2} \{ \text{Tr}(\mathbf{Y}^T \mathbf{M} \mathbf{Y}) \\ & + 2 \text{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{M} \mathbf{Y}) + \text{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{M} \mathbf{G} \mathbf{W}) \}, \end{aligned} \quad (10)$$

where the $M \times N$ matrix \mathbf{P} has the element $p^{old}(m|\mathbf{x}_n)$, the $M \times M$ matrix $\mathbf{M} = (\mathbf{I} - \mathbf{L})(\mathbf{I} - \mathbf{L})^T$, \mathbf{I} denotes the identity matrix, $d(\mathbf{v})$ denotes the diagonal matrix formed from the vector \mathbf{v} and $\mathbf{1}$ denotes the column vector of all ones. In order to obtain the weight matrix \mathbf{W} which minimizes the objective function (10), we take the derivative of (10) respect to \mathbf{W} and set it equal to zero, and then we obtain

$$\begin{aligned} \frac{\partial Q_{GLTP}}{\partial \mathbf{W}} = & -\mathbf{G} \mathbf{P} \mathbf{X} + \mathbf{G} d(\mathbf{P} \mathbf{1}) \mathbf{Y} + \mathbf{G} d(\mathbf{P} \mathbf{1}) \mathbf{G} \mathbf{W} \\ & + \sigma^2 \alpha \mathbf{G} \mathbf{W} + \sigma^2 \lambda \mathbf{G} \mathbf{M} \mathbf{Y} + \sigma^2 \lambda \mathbf{G} \mathbf{M} \mathbf{G} \mathbf{W}. \end{aligned} \quad (11)$$

We can find \mathbf{W} by solving a linear system:

$$[\mathbf{d}(\mathbf{P} \mathbf{1}) \mathbf{G} + \sigma^2 \alpha \mathbf{I} + \sigma^2 \lambda \mathbf{M} \mathbf{G}] \mathbf{W} = \mathbf{P} \mathbf{X} - (\mathbf{d}(\mathbf{P} \mathbf{1}) + \sigma^2 \lambda \mathbf{M}) \mathbf{Y} \quad (12)$$

Similarly we can obtain σ^2 by taking the corresponding derivative of (10) and set it equal to zero. We have

$$\begin{aligned} \sigma^2 = & \frac{1}{N_p D} (\text{Tr}(\mathbf{X}^T d(\mathbf{P}^T \mathbf{1}) \mathbf{X}) - 2 \text{tr}(\mathbf{Y}^T \mathbf{P} \mathbf{X})) \\ & - 2 \text{Tr}(\mathbf{W}^T \mathbf{G}^T \mathbf{P} \mathbf{X}) + \text{Tr}(\mathbf{Y}^T d(\mathbf{P} \mathbf{1}) \mathbf{Y}) \\ & + 2 \text{Tr}(\mathbf{W}^T \mathbf{G}^T d(\mathbf{P} \mathbf{1}) \mathbf{Y}) + \text{tr}(\mathbf{W}^T \mathbf{G}^T d(\mathbf{P} \mathbf{1}) \mathbf{G} \mathbf{W}). \end{aligned} \quad (13)$$

After the EM algorithm converges, the transformed point set is $\mathbf{Y}_{tran} = \mathbf{Y} + \mathbf{GW}$ and the probability of correspondence is stored in matrix \mathbf{P} . The pseudo-code of our proposed GLTP algorithm is given in Table 1.

Table 1. Pseudo-code for the proposed GLTP algorithm.

-
- Initialization: Initialize $0 \leq \omega \leq 1$, $K > 0$, $\alpha > 0$, $\beta > 0$, $\lambda > 0$ and $\sigma^2 = \frac{1}{DMN} \sum_{m,n=1}^{M,N} \| \mathbf{x}_n - \mathbf{y}_m \|^2$
 - Compute Gaussian kernel \mathbf{G} : $g_{ij} = \exp \frac{-1}{2} \left\| \frac{\mathbf{y}_i - \mathbf{y}_j}{\beta} \right\|^2$
 - Compute the LLE weight matrix \mathbf{L}
 - Compute matrix $\mathbf{M} = (\mathbf{I} - \mathbf{L})(\mathbf{I} - \mathbf{L})^T$
 - While (dissatisfy stopping criteria)
 - E-step:
Compute matrix \mathbf{P} according to (3)
 - M-step:
Compute weight matrix \mathbf{W} by solving (12)
Compute $N_p = \mathbf{1}^T \mathbf{P} \mathbf{1}$
Compute σ^2 according to (13)
Adjust α and λ by simulated annealing
 - End while
 - Obtain the transformed point set as $\mathbf{Y}_{tran} = \mathbf{Y} + \mathbf{GW}$
 - The probabilities of correspondence are stored in matrix \mathbf{P}
-

3.4. Parameter Selection

The proposed GLTP algorithm contains five free parameters: ω , K , β , α , and λ . ω ($0 \leq \omega \leq 1$) reflects the proportion of the points in \mathbf{X} which are treated outliers or noise, and β is the width of the Gaussian kernel. K is the number of neighbors used to compute the LLE coefficient matrix \mathbf{L} involved in (7) and (8), and it is related to the density and distribution of the points in the template \mathbf{Y} . Normally, the denser \mathbf{Y} is, the larger value K should be or vice versa. α and λ are the two trade-off parameters between the CPD and LLE regularization terms which also control the balance between the two topological constraints and the GMM-based matching term. In the GLTP algorithm, the two topological regularization terms play a complementary role to improve the robustness and accuracy of correspondence estimation at both global and local scales which is critical for non-rigid articulated registration. During EM iterations, the correspondence is largely determined in the initial stage which will not update significantly in the later iterations. To allow more accurate transformation estimation, it is helpful to slowly weaken the two topology constraints by reducing both α and λ via simulated annealing (SA). In this way, the GMM matching term becomes more and more dominant during the latter stage of EM optimization. This SA scheme is important to reduce the matching error by refining transformation estimation progressively.

4. Experiments

We implemented GLTP in Matlab which was tested on two 2D point sets used in [5, 6] (*fish* and *Chinese character*) and two 3D point sets used in [15] (*human face* and *bunny*). We also obtained a series of 3D human point sets of various poses from [1] and [17] as the test data along with a human template model (T-pose) from the MotionBuilder software. The human datasets create a challenging non-rigid highly articulated registration problem where different body shapes and highly articulated poses are involved. In the first experiment (2D point sets), we compare GLTP with CPD [15], TRS-RPM [6] and RPM-LNS [23] which were implemented using publicly available codes. Since RPM-LNS only handles 2D point sets and TRS-RPM did not perform well on the 3D data in our study, we only show the comparative analysis between GLTP and CPD in the last two experiments. In experiments below, we set $\beta = 2$, $\alpha = 3$, and $\omega = 0$ for the clean case or $\omega = 0.3$ for the cases with outliers or noise.

4.1. Results on 2D data

We consider five degrees of non-rigid deformations for each 2D data set. Fig. 2 shows the registration results of four methods. The registration error is computed as the average Euclidean distance between a point in the template and its counter part in the target based upon the ground truth correspondence. The four algorithms are evaluated by the mean and standard deviation (std) of the error over 100 trials at each level of deformation. GLTP ($\lambda = 1$, $K = 5$) provides accurate registration results for both data sets at different levels of deformations, and it has the best numerical results especially when the level of deformation increases.

We also study the cases with noise and outliers. Thanks to the consideration of outliers in the problem formulation (1), GLTP is comparable with CPD under outliers. However, in the case of noise, GLTP needs some adjustment regarding simulated annealing (SA) in the EM optimization (Table 1). Because the LLE term may be less effective under noise due to its reliance on the local structure which could be sensitive to noise, especially for sparse 2D point sets, we remove SA for the CPD term (i.e., α does not change) to impose the global coherence consistently. This yields a modified GLTP algorithm (mGLTP) that achieves similar robustness as CPD with respect to noise.

4.2. Results on 3D Data

For each of the two 3D point sets (*face* and *bunny*), we have a template point set and a target one. The two data sets are related through non-rigid transformation. We show the registration results in Fig. 3 for the clean case and the case with 50% outliers. It is shown that GLTP ($\lambda = 0.5$ and $K = 5$) performs well on both 3D data sets under both the clean and outlier cases.

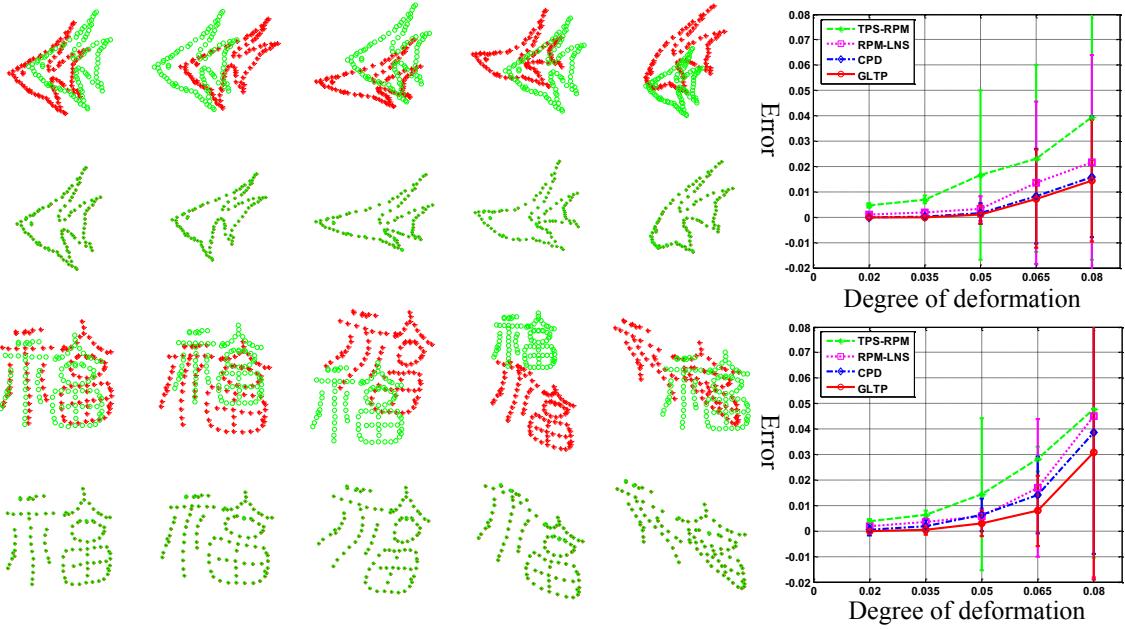


Figure 2. 2D point set registration results of the proposed GLTP algorithm on the *fish* (left top) and *Chinese character* (left bottom) with different levels of non-rigid deformations from left to right. For each group of experiments, the upper figures show the template (green) and target (red) models, and the lower figures are the registration results. The rightmost plots compare the four registration algorithms in terms of the mean and standard deviation of registration errors over 100 trials for different levels of deformations.

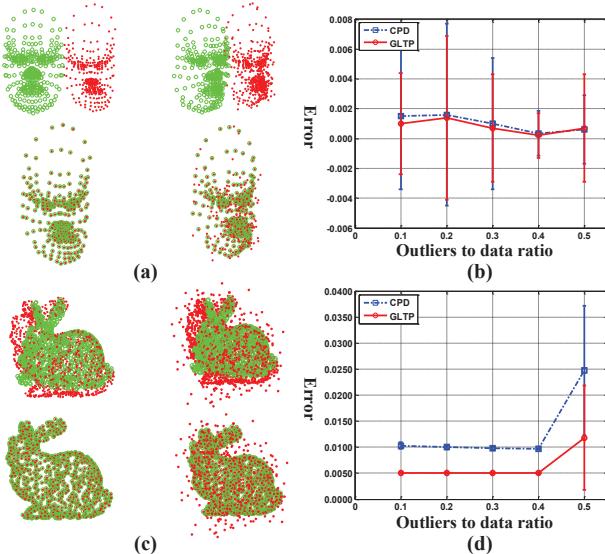


Figure 3. 3D point set registration results on *face* and *bunny*. In (a) and (c), the first row shows the template (green) and target (red) without/with outliers (left/right), and the second row gives the corresponding registration results without/with outliers (left/right). (b) and (d) compare GLTP and CPD in terms of the mean and standard deviation of registration errors over 20 trials for different ratios of outliers on *face* and *bunny*, respectively.

The numerical results under different outlier ratios are given in Fig. 3. For each level of outliers we tested GLTP and CPD 20 runs to compute the mean and std of the registration error. It is shown in Fig. 3 that GLTP is slightly better than CPD for the *face* and significantly better than CPD for the *bunny*. We believe it is because that the non-uniformity and relatively sparsity in the *face* data could weaken the local linear assumption and consequentially reduce the strength of the LLE-based topology constraint, while the *bunny* data are denser and more uniformly distributed which are beneficial for both LLE encoding and its influence on local matching. We also tested both 3D point sets under different noise levels. Similar to the 2D case, the mGLTP algorithm (with SA removed for the CPD term) can achieve similar performance as CPD.

4.3. Results on 3D Human Data

We use the labeled T-pose human template (643 points, Fig. 1 (column (a) and first row)) and five human point sets (12500 points) captured by 3D laser scanner from [1] as the target data. There are significant shape differences and non-rigid articulated deformations between the template and the target data. Therefore we set $\alpha = 10$, $\lambda = 5 \times 10^6$ and $K = 8$ to strengthen the LLE constraint and to partially offset the CPD constraint that tends to discourage non-coherent local deformation. We first compare GLTP with CPD on the clean data in Fig. 4. When articulated defor-

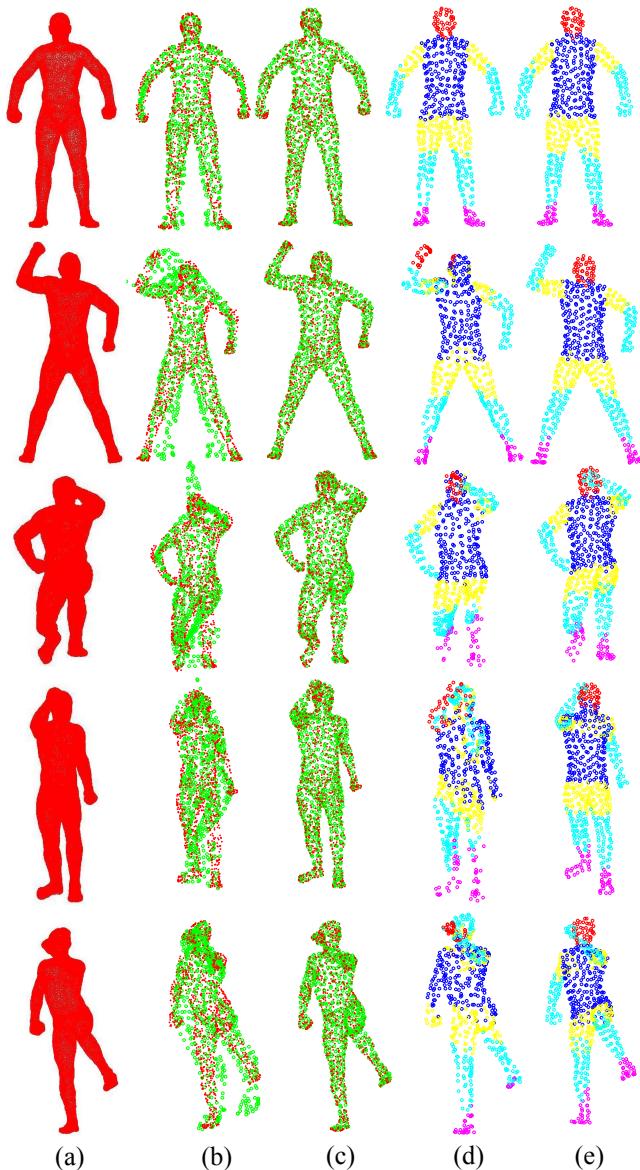


Figure 4. Registration results for the 3D human data. Column (a) shows the five target models with five different poses; columns (b) and (c) show the registration results of CPD and GLTP respectively; columns (d) and (e) show the correspondence estimation results of CPD and GLTP respectively (color display needed).

mation is not significant between the template and the target, such as the first pose, CPD can give correct correspondence estimation and its registration result is close to that of GLTP. However, in the cases of highly articulated deformations, e.g., poses 2 to 5, significant correspondence errors are observed from the CPD results around the head, limbs and body joints, leading to large registration errors in those areas. On the other hand, GLTP provides stable and accurate correspondence estimation across all poses which contributes to good whole-body registration results.

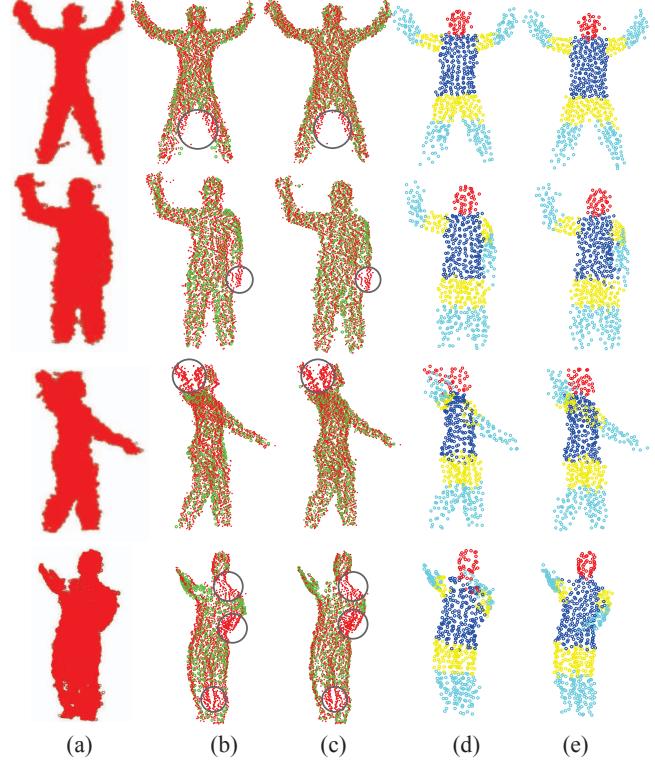


Figure 5. Registration results for the 3D human point sets created from the depth data. Column (a) shows the four target models with different poses; columns (b) and (c) show the registration results of CPD and GLTP respectively; columns (d) and (e) show the correspondence estimation results of CPD and GLTP respectively (color display needed).

Furthermore, we compare the two algorithms under the point sets obtained from the Kinect depth sensor. We use the same labeled T-pose human template and the four human point sets (around 20000 points) created by the two-view depth maps from [17] as the target data. Since the depth data are noisy from Kinect, we set $\omega = 0.3$ to accommodate stronger noise. The comparative results are shown in Fig. 5. It is observed that CPD and GLTP yield similar results when the articulated deformation is not significant between the template and the target, such as the 1st and 2nd poses. However, in the cases of highly articulated deformations, e.g., the 3rd and 4th poses, CPD leads to obvious correspondence errors (around the right hand and the head for the 3rd pose and around the left arm and the torso for the 4th pose). On the other hand, GLTP provides stable and more accurate correspondence estimation across all poses. Moreover due to stable and accurate correspondence estimation across all poses, GLTP provides better registration results (see circled areas in Fig. 5 columns (b) and (c)) than CPD around the whole body. This result will allow accurate pose estimation that will be our future research focus.

4.4. Discussion

Although GLTP is promising in handling complicated non-rigid transformation including highly-articulated deformation, it has two limitations with respect to the LLE-based topological constraint that are worth further investigation. First, the strength of the LLE term is compromised under noise due to its dependency on local neighborhood structures in the target data that are sensitive to noise. mGLTP slightly mitigates this problem by adjusting the SA scheme of the CPD term at the price of the flexibility to handle articulated deformation. A possible remedy is to denoise point sets (e.g., [8]) prior to registration. Second, LLE is more reliable when the point sets are more uniform and denser. In the area of sparse or non-uniform points, LLE encoding may involve, for a given reference point, some neighbors which are from different local segments with different articulated deformations, resulting unrealistic local topology. It is possible to apply dense interpolation of points (e.g., [9]) or use a “neighborhood ball” to choose a variable number of neighbors for each point which are likely to be on the same body part.

5. Conclusion

We have presented a new GMM-based point set registration algorithm, GLTP, which is aimed at complex non-rigid transformation including highly articulated deformation. Unlike most algorithms which make some assumptions (e.g., locally rigid, no significant deformation between two data sets or using a large training data set for pose initialization) or require certain initial conditions (e.g., pre-segmentation, part-based representation, correspondence initialization) to attack this problem, GLTP is more general and flexible. The key idea of GLTP is the incorporation of the two topologically complementary (global/local) constraints into a unified probability density estimation framework. The experiment results on the classic 2D and 3D data show that GLTP outperforms three recent methods for various deformations with good robustness under outliers and noise. In particularly, the experiments on a set of 3D human point data further manifest the advantages of our algorithm for handling highly articulated and non-rigid deformation, even under noise or outliers. Our future work will be focused on the two limitations as well as the further application of GLTP to the real-world depth data.

References

- [1] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. Scape: Shape completion and animation of people. *ACM Trans. Graph.*, 24:408–416, 2005.
- [2] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *TPAMI*, 24(4):509–522, 2002.
- [3] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *TPAMI*, 14(2):239–256, 1992.
- [4] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, Inc., New York, NY, USA, 1995.
- [5] H. Chui and A. Rangarajan. A Feature Registration Framework using Mixture Models. In *Proc. IEEE MMBIA Workshop*, 2000.
- [6] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *CVIU*, 89(2-3):114–141, 2003.
- [7] R. Horaud, F. Forbes, M. Yguel, G. Dewaele, and J. Zhang. Rigid and articulated point registration with expectation conditional maximization. *TPAMI*, 33(3):587–602, 2011.
- [8] B. Huhle, T. Schairer, P. Jenke, and W. Strasser. Robust non-local denoising of colored depth data. In *Proc. CVPR Workshop*, 2008.
- [9] Z. Jia, Y.-J. Chang, T.-H. Lin, and T. Chen. Dense interpolation of 3D points based on surface and color. In *Proc. ICIP*, 2011.
- [10] B. Jian and B. C. Vemuri. Robust point set registration using Gaussian mixture models. *TPAMI*, 33(8):1633–1645, 2011.
- [11] B. Jian, B. C. Vemuri, and B. C. Vemuri. A robust algorithm for point set registration using mixture of Gaussians. In *Proc. ICCV*, 2005.
- [12] Y. T. Li X, P. TE, G. JC, and D. BM. Constrained non-rigid registration for whole body image registration: Method and validation. *Proc. SPIE Medical Imaging: Image Processing*, 6512, 2007.
- [13] J. Ma, J. Zhao, J. Tian, Z. Tu, and A. L. Yuille. Robust estimation of nonrigid transformation for point set registration. In *Proc. CVPR*, 2013.
- [14] D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer. Articulated shape matching using laplacian eigenfunctions and unsupervised point registration. In *Proc. CVPR*, 2008.
- [15] A. Myronenko and X. Song. Point set registration: Coherent point drift. *TPAMI*, 32(12):2262–2275, 2010.
- [16] A. Myronenko, X. Song, and M. A. Carreira-Perpinan. Non-rigid point set registration: Coherent point drift (CPD). In *Proc. NIPS*, 2006.
- [17] F. Ofli, R. Chaudhry, G. Kurillo, R. Vidal, and R. Bajcsy. Berkeley MHAD: A comprehensive multimodal human action database. In *Proc. WACV*, 2013.
- [18] S. Pellegrini, K. Schindler, , and D. Nardi. A generalization of the ICP algorithm for articulated bodies. In *Proc. BMVC*, 2008.
- [19] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000.
- [20] A. Weiss, D. Hirshberg, and M. J. Black. Home 3D body scans from noisy image and range data. In *Proc. ICCV*, 2011.
- [21] M. Ye, X. Wang, R. Yang, L. Ren, and M. Pollefeys. Accurate 3d pose estimation from a single depth image. In *Proc. ICCV*, pages 731–738, 2011.
- [22] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *IJCV*, 13(2):119–152, 1994.
- [23] Y. Zheng and D. Doermann. Robust point matching for nonrigid shapes by preserving local neighborhood structures. *TPAMI*, 28(4):643–649, 2006.