Non-rigid Articulated Point Set Registration with Local Structure Preservation

Song Ge and Guoliang Fan*
School of Electrical and Computer Engineering
Oklahoma State University, Stillwater, OK 74078
{song.ge,guoliang.fan}@okstate.edu

Abstract

We propose a new Gaussian mixture model (GMM)-based probabilistic point set registration method, called Local Structure Preservation (LSP), which is aimed at complex non-rigid and articulated deformations. LSP integrates two complementary shape descriptors to preserve the local structure. The first one is the Local Linear Embedding (LLE)-based topology constraint to retain the local neighborhood relationship, and the other is the Laplacian Coordinate (LC)-based energy to encode the local neighborhood scale. The registration is formulated as a density estimation problem where the LLE and LC terms are embedded in the GMM-based Coherent Point Drift (CPD) framework. A closed form solution is solved by an Expectation Maximization (EM) algorithm where the two local terms are jointly optimized along with the CPD coherence constraint. The experimental results on a challenging 3D human dataset show the accuracy and efficiency of our proposed approach to handle non-rigid highly articulated deformations.

1. Introduction

Point set registration is a fundamental issue in computer vision. The registration techniques usually fall into two categories: rigid and non-rigid depending on the underlying transformation model. Iterative Closest Point (ICP) [2, 26] is a classic rigid registration method which iteratively assigns correspondence and then finds the least square transformation by using the estimated correspondence. For non-rigid registration, shape features are commonly used for correspondence initialization [27, 11, 12] or directly involved in the matching process [22, 20]. Gaussian Mixture Model (GMM)-based registration algorithms are becoming an important category, where the point sets are represented by density functions and then registration is cast as either a density estimation problem [4, 5, 16, 15] or an alignment between two distributions [8, 9]. Coherent Point Drift (CPD) [16, 15] is an effective GMM-based non-rigid registration algorithm that imposes a coherence constraint to preserve the topology of the point sets.

On the other hand, articulated point set registration is becoming an attractive topic, e.g, human pose estimation [23] or body shape modeling [21]. Some initial conditions or new constraints are needed to cope with the complicated non-rigid articulated deformation. In [23], a pose initialization from a large training dataset is used to reduce the articulated deformation between two point sets to be registered. In [13], Laplacian eigenfunctions are used to align two voxels represented by spectral graphs for registration initialization. By incorporating the global (CPD) and local topological constraints (LLE) into a GMM-based framework, the Global-Local Topology Preservation (GLTP) algorithm was proposed in [6] which obtains promising performance. Some approaches treat the articulated deformation as a chain of constrained rigid motions, where pre-segmentation and locally rigid assumption are usually required. An articulated ICP (AICP) algorithm was proposed in [18] which adopts a divide-and-conquer strategy to iteratively estimate the articulated structure by assuming it is partially rigid. The similar strategy was applied in [7] by using a GMM-based rigid registration algorithm instead of ICP. By using the exponential maps based parametrization, the articulated deformation is effectively embedded in the GMM-based framework in [24].

We propose a new GMM-based point set registration algorithm, called *Local Structure Preservation* (LSP), to deal with non-rigid highly articulated deformations, such as 3D human data. Without involving the locally rigid assumption or requiring any initial conditions, LSP is non-rigid both globally and locally which is more flexible and general to handle the non-rigid articulated deformations than those with the locally rigid assumption (e.g. [18, 7, 24]). The key is to involve two local shape descriptors, the Local Linear Embedding (LE) and Laplacian Coordinate (LC), to preserve the local structure in terms of the *neighborhood relationship* and the *neighborhood scale*, respectively. The

*This work is supported by the Oklahoma Center for the Advancement of Science and Technology (OCAST) under grant HR12-30 and the National Science Foundation (NSF) under grant NRI-1427345.

two local regularization terms are unified with the CPD coherence constraint into the GMM-based density estimation framework. A joint optimization of three terms is achieved by an Expectation and Maximization (EM) algorithm for Maximum Likelihood (ML) optimization.

2. Proposed Method

We first review the GMM-based registration framework along some existing regularization terms for non-rigid transformation estimation. Then we will introduce a new local regularization term E_{LSP}. Afterwards, we present the formulation and optimization of the proposed LSP algorithm followed by a discussion of parameter selection.

2.1. Registration as Mixture Density Estimation

We consider point set registration as a probability density estimation problem, where one point set ($\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_M]^T$, $\mathbf{y}_m \in R^D$) is treated as the template with a sparse distribution which presents the centroids from the Gaussian mixture model (GMM), while the other point set ($\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N]^T$, $\mathbf{x}_n \in R^D$) is served as the target with a dense distribution, and M, N, D are the number of points and the dimension of each point respectively. Then the goal is to find the optimal GMM parameters (e.g. the centroids which are controlled by unknown transformation) to maximize the GMM posterior probability. We denote the spatial transformation as $\mathcal{T}(\mathbf{Y}, \mathbf{\Theta})$, which can be considered as a function of $\mathbf{Y}$ with parameters $\mathbf{\Theta}$. To account for outliers in $\mathbf{X}$, a uniform component with weight ω ($0 \leq \omega \leq 1$) is added [15, 4]. For simplicity, we consider all Gaussian components are independently distributed with an equal isotropic variance σ^2 and an equal weight, then the joint GMM probability density function is written as

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

where $p(\mathbf{x}|m) = \frac{1}{(2\pi\sigma^2)^{\frac{D}{2}}} \exp(-\frac{\|\mathbf{x}-\mathcal{T}(\mathbf{y}_m,\mathbf{\Theta})\|^2}{2\sigma^2})$ ($m = 1, \cdots, M$), $p(\mathbf{x}|m) = \frac{1}{N}$ for $m = M+1$, $\pi_m = \frac{1-\omega}{M}$ ($m = 1, \cdots, M$), and $\pi_{M+1} = \omega$. Then registration is converted to the problem of finding $\mathbf{\Theta}$ and σ^2 that minimizes the negative log-likelihood of (1).

Following the EM algorithm for GMM-based clustering [3, 15], we can find the objective (E-step) as

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

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

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

Then in the M-step the new GMM parameters are updated by minimizing (2). The EM algorithm performs iteratively by alternating between E-step and M-step until it converges. The different type of transformations can lead to different optimization strategies.

2.2. Non-rigid Transformation Regularization

Particularly, Coherent Point Drift (CPD) [15] is a powerful and noteworthy GMM-based non-rigid registration method where the underlying non-rigid transformation $\mathcal{T}(\mathbf{Y}, \mathbf{\Theta})$ is defined as the initial position $\mathbf{Y}$ plus a displacement function $f(\mathbf{Y})$. The displacement function f is modeled in a Reproducing Kernel Hilbert Space (RKHS) and the spatial smoothness regularization is defined as the Fourier domain norm of f. It has been proved that the optimal function f which minimizes the objective (2) under the spatial smoothness constraint is given by a linear combination of Gaussian kernel functions as

$$\mathcal{T}(\mathbf{Y}, \mathbf{W}) = \mathbf{Y} + \mathbf{G}\mathbf{W}, \qquad (4)$$

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. The corresponding constraint which regularizes the weight matrix $\mathbf{W}$ has the form as

$$E_{MC}(\mathbf{W}) = \mathrm{Tr}(\mathbf{W}^T \mathbf{G} \mathbf{W}). \qquad (5)$$

It was shown in [15] that the selection of the Gaussian kernel makes the regularization equivalent to the one in the Motion Coherence Theory [25] which forces points to move together as a group to keep the motion coherence. The motion coherence is helpful to keep the overall spatial connectivity of a complicated point set during the registration. However, it may not handle the non-rigid highly articulated deformation where the motion coherence assumption may be violated.

Recently some specific regularization terms were proposed and embedded in the GMM-based CPD framework for improving registration performance. In [17], a Graph-Laplacian (GL) based regularization term, which enforces transformation smoothness to avoid mis-matches during registration, is added to restrict a large displacement between two neighboring points. This regularization term enhances the registration robustness under noise, outliers and occlusions. It mainly penalizes large displacements between the adjacent neighbors, but does not consider the spatial relationship between the edges (constructed by neighboring points) which may not be effective to preserve the local shape structure. On the other hand, the LLE-based topology constraint proposed in [6] is intended to preserve the neighborhood structure by retaining the local neighborhood relationship. This was found helpful to cope with non-rigid articulated deformations. However, the LLE-based neighborhood structure is insensitive to the scaling which means the scale of the neighborhood structure may not be preserved well, leading to inaccurate correspondence estimation at body joints, where two adjacent segments are connected. However, joint estimation is critical for articulated pose estimation. This motivates us to study more effective regularization for local structure preservation.

2.3. Local Structure Preservation

In this work, we try to preserve the local shape structure in terms of neighborhood structure along with the motion coherence with respect to two aspects: spatial relative location and the size of the neighborhood structure. Our proposed LSP regularization term E_{LSP} is built for preserving two local shape descriptors. One is the LLE-based descriptor in terms of preserving the spatial relationship among the neighboring points. The other one is a Laplacian coordinate (LC)-based descriptor to preserve the scale of the neighborhood structure which may not be encoded by the LLE-based descriptor. By preserving two local shape descriptors, E_{LSP} has the form as

$$E_{LSP} = E_{LL} + E_{LC}, \tag{6}$$

where E_{LL} and E_{LC} denote the LLE-based and LC-based regularization terms for local structure preservation respectively.

For E_{LL} we adopt the similar idea of GLTP [6]. By finding the K_1 nearest neighbors of each point in $\mathbf{Y}$ according the Euclidean distance and representing each point in $\mathbf{Y}$ by a weighted linear combination of its neighbors, the LLE-based regularization term in matrix form is

$$E_{LL}(\mathbf{W}) = \\ \mathrm{Tr}(\mathbf{Y}^T\mathbf{M}\mathbf{Y}) + 2\,\mathrm{Tr}(\mathbf{W}^T\mathbf{G}\mathbf{M}\mathbf{Y}) + \mathrm{Tr}(\mathbf{W}^T\mathbf{G}\mathbf{M}\mathbf{G}\mathbf{W}), \tag{7}$$

where the $M\times K$ matrix $\mathbf{C}$ contains the neighborhood information for each point in $\mathbf{Y}$, the $M\times M$ matrix $\mathbf{M} = (\mathbf{I}-\widehat{\mathbf{C}})(\mathbf{I}-\widehat{\mathbf{C}})^T$, $\widehat{\mathbf{C}}$ is an $M\times M$ matrix with $\widehat{\mathbf{C}}_{ij} = \mathbf{C}_{ij}$, for $j \le K$ and $\widehat{\mathbf{C}}_{ij} = 0$ for $j > K$, $\mathbf{I}$ denotes the identity matrix. And the objective is to find the optimal coefficient matrix $\mathbf{W}$ which controls the non-rigid transformation to minimize (7).

The Laplacian coordinate is a simple form of the differential coordinate which has been used to describe the local shape structure in mesh deformation [19, 10]. Given a mesh model $(\mathbf{V}, \mathbf{E})$, where $\mathbf{V}$ denotes the set of vertices and $\mathbf{E}$ denotes the edge set, the Laplacian coordinate could be expressed as

$$\mathcal{L}(\mathbf{v}_i) = \sum_{(i,j)\in\mathbf{E}} \mathbf{A}_{ij}(\mathbf{v}_i - \mathbf{v}_j), \tag{8}$$

where $\mathbf{A}_{ij}$ is the weight coefficient and (i, j) denotes the edge constructed by vertices $\mathbf{v}_i$ and $\mathbf{v}_j$. It is easy to show that the Laplacian coordinate is sensitive with scaling which is helpful to preserve the size of neighborhood structure. There are two steps to model the LC-based regularization in terms of point set registration. First, find the K_2 nearest neighbors of each point in $\mathbf{Y}$ according the Euclidean distance which is used to model the graph. Then the regularization term which measures the difference of the Laplacian coordinate after the non-rigid transformation has the form as

$$E_{LC}(\mathbf{W}) = \sum_{m=1}^{M} \|\mathcal{L}(\mathbf{y}_m) - \mathcal{L}(\mathbf{y}_m + \mathbf{G}(m,\cdot)\mathbf{W})\|^2, \tag{9}$$

where $\mathbf{G}(m,\cdot)$ denotes the mth row of $\mathbf{G}$. We can rewrite it in the matrix form as

$$E_{LC}(\mathbf{W}) = \mathrm{Tr}(\mathbf{W}^T\mathbf{G}\mathbf{L}^T\mathbf{L}\mathbf{G}\mathbf{W}), \tag{10}$$

where the $M\times M$ Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{A}$, matrix $\mathbf{A}$ is the adjacency matrix of the graph, $\mathbf{D}$ is the degree matrix which is a diagonal matrix and its entries are column sums of $\mathbf{A}$ [14]. For simplicity we set $\mathbf{A}_{ij} = 1$ if points $\mathbf{y}_i$ and $\mathbf{y}_j$ are neighbors and $\mathbf{A}_{ij} = 0$ otherwise. Now the objective is to find the optimal weight matrix $\mathbf{W}$ to preserve the neighborhood structure by minimizing (10). We need to point out the Laplacian coordinate is not rotation-invariant, therefore this strong constraint may be counter-productive for the articulated deformation. However a balanced control of E_{LL} and E_{LC} under CPD-based regularization makes them complement each other to preserve the local structure.

2.4. Algorithm and Optimization

In this work we treat point set registration problem as a GMM density estimation problem. By incorporating the motion coherence and the local structure preservation regularization into general formulation (2), the registration objective function can be written in the form as

$$Q(\mathbf{W},\sigma^2) = \sum_{m,n=1}^{M,N} p^{old}(m|\mathbf{x}_n)\frac{\|\,\mathbf{x}_n - (\mathbf{y}_m + \mathbf{G}(m,\cdot)\mathbf{W})\,\|^2}{2\sigma^2} \\ + \frac{N_pD}{2}ln(\sigma^2) + \frac{\alpha}{2}E_{MC}(\mathbf{W}) + \frac{\lambda}{2}E_{LL}(\mathbf{W}) + \frac{\gamma}{2}E_{LC}(\mathbf{W}), \tag{11}$$

where α, λ and γ are the trade-off parameters among the GMM matching term and the regularization terms. Then we can rewrite the objective function (11) in a matrix form. In order to obtain the weight matrix $\mathbf{W}$ which minimizes the objective function, we take the derivative of the matrix form of (11) respect to $\mathbf{W}$ and set it equal to zero, and then we obtain

$$\frac{\partial Q}{\partial \mathbf{W}} = -\mathbf{GPX} + \mathbf{G}d(\mathbf{P1})\mathbf{Y} + \mathbf{G}d(\mathbf{P1})\mathbf{GW} + \sigma^2\alpha\mathbf{GW} + \sigma^2\lambda\mathbf{GMY} + \sigma^2\lambda\mathbf{GMGW} + \sigma^2\gamma\mathbf{GL}^T\mathbf{LGW} = 0, \quad (12)$$

where the $M \times N$ matrix $\mathbf{P}$ has the element $p^{old}(m|\mathbf{x}_n)$, $d(\mathbf{v})$ denotes the diagonal matrix formed from the vector $\mathbf{v}$ and $\mathbf{1}$ denotes the column vector of all ones. Then $\mathbf{W}$ can be obtained by solving a linear system:

$$[d(\mathbf{P1})\mathbf{G} + \sigma^2\alpha\mathbf{I} + \sigma^2\lambda\mathbf{MG} + \sigma^2\gamma\mathbf{L}^T\mathbf{LG}]\mathbf{W} = \mathbf{PX} - (d(\mathbf{P1}) + \sigma^2\lambda\mathbf{M})\mathbf{Y}. \quad (13)$$

Similarly σ^2 can be obtained by taking the corresponding derivative of the matrix form of (11) and setting it equal to zero. We have

$$\sigma^2 = \frac{1}{N_p D}(\mathrm{Tr}(\mathbf{X}^T d(\mathbf{P}^T\mathbf{1})\mathbf{X}) - 2tr(\mathbf{Y}^T\mathbf{PX}) - 2\,\mathrm{Tr}(\mathbf{W}^T\mathbf{G}^T\mathbf{PX}) + \mathrm{Tr}(\mathbf{Y}^T d(\mathbf{P1})\mathbf{Y}) + 2\,\mathrm{Tr}(\mathbf{W}^T\mathbf{G}^T d(\mathbf{P1})\mathbf{Y}) + tr(\mathbf{W}^T\mathbf{G}^T d(\mathbf{P1})\mathbf{GW})). \quad (14)$$

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 EM optimization process is similar with that in [15, 6].

2.5. Parameter Discussion

The proposed algorithm contains several free parameters: ω, K_1, K_2, β, α, λ and γ. ω $(0 \leq \omega \leq 1)$ reflects the proportion of the points in $\mathbf{X}$ which are treated as outliers or noise, and β is the width of the Gaussian kernel. K_1 is the number of neighbors used to compute the LLE coefficient matrix $\mathbf{C}$ and K_2 is the number of neighbors to compute the Laplacian matrix $\mathbf{L}$. The values of K_1 and K_2 are related to the density and distribution of the points in the template $\mathbf{Y}$. Normally, the denser $\mathbf{Y}$ is, the larger values should be or vice versa. α, λ and γ are the three trade-off parameters among the motion coherence and the local structure preservation terms, which also control the balance between the regularization terms and the GMM-based matching term.

In the proposed LSP algorithm, the three regularization terms are incorporated together to play a mutually helpful role to preserve the local structure along with the spatial coherence, which is important to obtain the accurate correspondence estimation for complicated non-rigid articulated deformations. During the EM optimization process, the correspondence estimation during the initial iterations is critical to guide the template point set to move to the target along the desirable directions. Therefore, we set a relative large value for λ comparing with α and γ in the initial stage to strengthen the local structure preservation under motion coherence. However the strong regularization may discourage the registration to conform to local details. So during the latter iterations, we slowly reduce α, λ and γ to allow the points moving close to the corresponding points. In particular, for a clean target point set, we finally decrease the α and γ down to zero and only preserve the LLE-based local structure to provide good registration results with local detail shape representation.

3. Experiments

We have implemented the proposed algorithm in Matlab and tested it on a 3D human mesh dataset which creates a challenging non-rigid articulated registration problem where highly articulated poses are involved. The target point sets (12500 points for each target) are extracted from the SCAPE human mesh dataset [1] which includes 35 different highly articulated poses and are fully registered. The template point set we use is a T-pose model (2150 points) provided by [24] which has significant shape differences and articulated deformations compared with the target point sets.

3.1. Data Preparation

Since the template and target models are captured from different subjects and also have different number of points, it is difficult to obtain the ground-truth correspondences. Thus a quantitative result in terms of registration error is not available in this experiment. Instead we use the accuracy of body segment labeling to evaluate the registration performance. To do so, we need to create the ground-truth segment labels for the template model and all target models. The datasets provide the ground-truth joint positions both for the "T-pose" template model and one target model, from which we can create the segment labels. The general T-pose template model with joint positions and segment labels is shown in Fig. 1 and the target model with joint positions and segment labels, called "initial target pose" is shown in Fig. 2(a)(b).

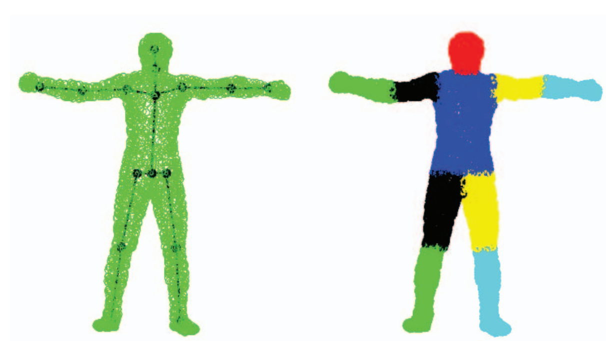


Figure 1. The T-pose template model with given joint positions (left) and segment labels (right).

Since the SCAPE mesh data are fully aligned and registered, the body segment labels can be easily transferred from one labeled target model to all others. One example is shown in Fig. 2 where the segment labeling information from one labeled target model is transferred to another target model of an arbitrary pose, as shown Fig. 2(c)(d). With the creation of ground-truth segment labels for all target models, we can evaluate the registration performance quantitatively by calculating the labeling accuracy of each body segment after registration. For each point in the template model, we propagate its segment label to the corresponding point in the target model by estimated correspondence. If this assigned segment label is as the same as the ground-truth label, we treat it as the correct segment label as shown in Fig. 3. Then the labeling accuracy is calculated as the percentage of the points with correct segment labels over all labeled points. It is worth mentioning that the labeling information is not involved during the registration process and is only for performance evaluation.

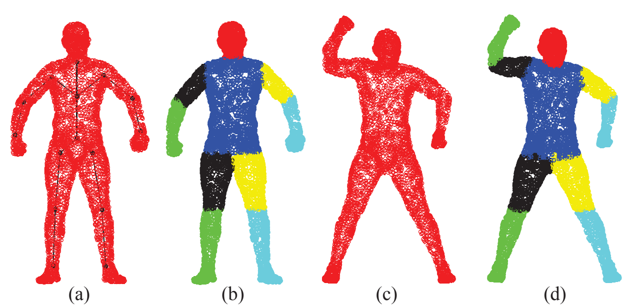


Figure 2. The creation of ground-truth segment labels. (a) The "initial pose" target model with given joints positions; (b) The labeled target model from (a); (c) A target model of an arbitrary pose; (d) The labeled target model obtained by label transferring from (b).

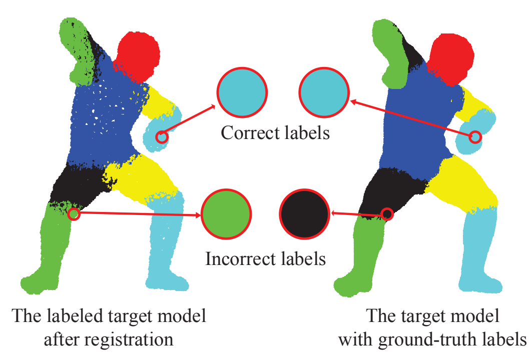


Figure 3. The illustration of the calculation of labeling accuracy after registration where all labeled points in the registered target model (left) are checked with respect to the ground truth labels created for the target model according to Fig. 2 (right).

3.2. Results on 3D Human Data

In order to evaluate the performance of our proposed algorithm, we performed a quantitative comparison against several recent algorithms which are able to handle 3D data. To the best of our knowledge, there is little study and numerical results on 3D non-rigid point set registration, especially for those with complicated articulated deformations. Therefore, we mainly compared our proposed algorithm LSP with CPD [15], GLTP [6], and the algorithm proposed in [17] (for simplicity we refer it as GL-CPD and for fair comparison we implemented it without feature-based correspondence initialization). In the experiments, we set $\alpha = 40, \lambda = 5 \times 10^6$, $\beta = \sqrt{2}$, $\gamma = 10$, $\omega = 0$ and $K_1 = K_2 = 15$. The extra small or large values of K_1 and K_2 cannot describe a good local neighborhood structure. In the experiment, we found the optimal values of K_1 and K_2 are between 10 to 20 and the registration results are not very sensitive with the values in that range.

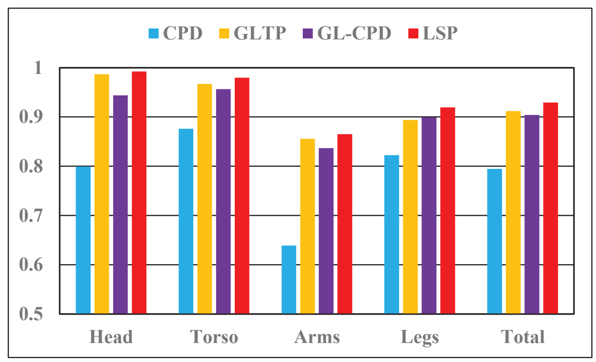
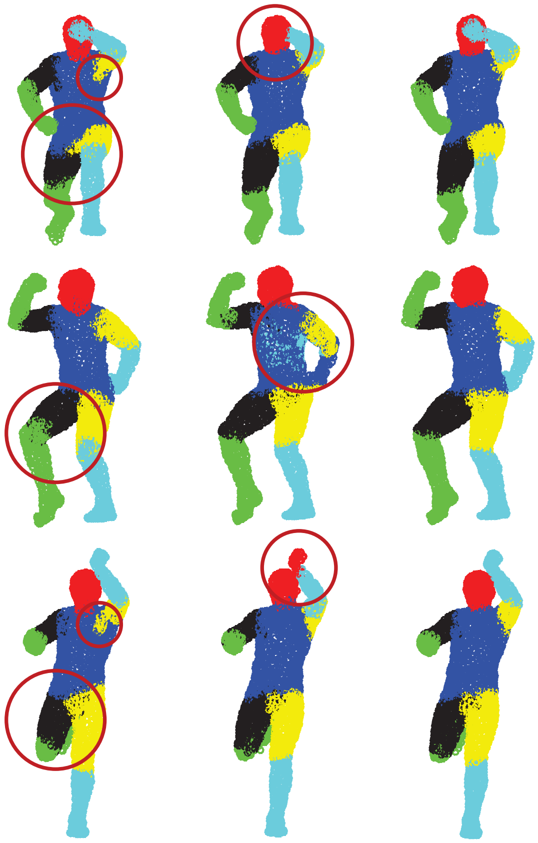


Figure 5. Quantitative comparison of CPD, GLTP, GL-CPD and LSP for the SCAPE data in terms of the accuracy of segment labeling at some body segments and the whole body.

We first present the qualitative comparison in Fig. 4. It is shown that when articulated deformation is not significant between the template and the target, such as the first and second poses, all four algorithms (CPD, GLTP, GL-CPD and LSP) can obtain the similar results of correspondence estimation (segment labels shown in columns (b-e)). However, in the cases of highly articulated deformations, e.g., poses 3 to 6, significant correspondence errors are observed from the CPD and GL-CPD results around the head, limbs and body joints, leading to large registration errors in those areas. GL-CPD is helpful to improve the results compared with CPD, but it is not sufficient enough to keep the local structure in some cases (e.g. pose 4 to 6). GLTP provides better correspondence estimation across all poses at the global scale. But at the local scale, the correspondence estimation at body joints, the connection part between adjacent segments, is not very accurate due to the scaling in-

Figure 4. Qualitative comparison of CPD, GLTP, GL-CPD and LSP for the SCAPE data. Column (a) shows the six target models with different poses (from the first to the sixth rows); columns (b-e) show the correspondence estimation results of CPD, GLTP, GL-CPD and LSP respectively (the color labels are transferred from the template model through the estimated correspondence); columns (f-i) show the registration results of CPD, GLTP, GL-CPD, and LSP respectively.



(a) (b) (c)

Figure 6. Detailed comparison of the segment labeling of GLTP (column (a)), GL-CPD (column (b)) and LSP (column (c)) at the arms, torso and legs (the circled areas).

sensitivity of the LLE-based local shape descriptor. On the other hand, LSP provides stable and accurate correspondence estimation across all poses which directly contributes to good registration results around the whole body and at local segments. The improvements in the joint areas are obvious, compared with other three algorithms.

As an alternative of correspondence evaluation, the mean labeling accuracy of body segments over all 35 target models is compared among CPD, GLTP, GL-CPD and LSP in Fig. 5. From Fig. 5 we can see that GLTP, GL-CPD and LSP are significantly better than CPD, which shows the advantages by incorporating additional regularization terms for dealing with non-rigid articulated registrations. For the arms and torso, GLTP achieves better results than GL-CPD, and both GLTP and GL-CPD have similar results on the leg segments. The main reason that GLTP outperforms GL-CPD at the arms and torso is due to the fact that large articulated deformations may not be represented well by the graph-based distance constraint in GL-CPD. On the other hand, GL-CPD better preserves the segment length than GLTP, especially near the knees on the two legs, leading a better labeling accuracy at the legs. LSP outperforms the other three algorithms at all segments. The quantitative results illustrate the advantage of the joint use of two local shape descriptors to handle the complicated non-rigid articulated deformation. More detailed comparison results between GLTP, GL-CPD and LSP in terms of correspondences estimation at local segments are shown in Fig. 6.

4. Conclusion

We have presented a new GMM-based point set registration algorithm, referred to as *Local Structure Preservation* (LSP), which is targeted on complex non-rigid and articulated deformations. The proposed LSP algorithm embraces the original CPD constraint and more importantly, introduces dual local regularization terms to preserve the local structure for non-rigid registration. One is the LLE-based shape descriptor and the other is the LC-based neighborhood structure. The joint use of them in the CPD framework demonstrates better flexibility and capability of coping with non-rigid and articulated deformations while retaining the size of each segment and spatial conference. Compared with the several related algorithms, including CPD, GLTP and GL-CPD, LSP is able to provide more accurate registration results around joints where two adjacent segments are connected. This is especially important in the case of articulated pose estimation. The experiment results on a set of 3D human data manifest the advantages of our LSP algorithm for handling non-rigid and highly articulated body deformations. Our future research will be extended to depth-based articulated pose estimation as well as the registration of non-rigid and articulated 2D/3D objects.

References

[1] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rodgers, and J. Davis. SCAPE: Shape completion and animation of people. *ACM Transactions on Graphics*, 24:408–416, 2005.

[2] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[3] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, Inc., New York, NY, USA, 1995.

[4] H. Chui and A. Rangarajan. A feature registration framework using mixture models. In *Proceedings of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 190–197, 2000.

[5] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3):114–141, 2003.

[6] S. Ge, G. Fan, and M. Ding. Non-rigid point set registration with global-local topology preservation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 245–251, 2014.

[7] R. Horaud, F. Forbes, M. Yguel, G. Dewaele, and J. Zhang. Rigid and articulated point registration with expectation conditional maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):587–602, 2011.

[8] B. Jian and B. C. Vemuri. A robust algorithm for point set registration using mixture of Gaussians. In *Proceedings of IEEE International Conference on Computer Vision*, pages 1246–1251, 2005.

[9] B. Jian and B. C. Vemuri. Robust point set registration using Gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–45, 2011.

[10] Y. Lipman, O. Sorkine, D. Cohen-Or, D. Levin, C. Rossi, and H.-P. Seidel. Differential coordinates for interactive mesh editing. In *Proceedings of Shape Modeling Applications*, pages 181–190, 2004.

[11] J. Ma, J. Zhao, J. Tian, Z. Tu, and A. L. Yuille. Robust esitimation of nonrigid transformation for point set registration. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 2147–2154, 2013.

[12] J. Ma, J. Zhao, J. W. Tian, A. L. Yuille, and Z. W. Tu. Robust point matching via vector field consensus. *IEEE Transactions on Image Processing*, 23(4):1706–1721, 2014.

[13] D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer. Articulated shape matching using Laplacian eigenfunctions and unsupervised point registration. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[14] R. Merris. Laplacian matrices of graphs: a survey. *Linear Algebra and its Applications*, 197-198(0):143–176, 1994.

[15] A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 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 *Proceedings of Advances in Neural Information Processing Systems*, pages 1009–1016, 2006.

[17] V. Panaganti and R. Aravind. Robust nonrigid point set registration using Graph-Laplacian regularization. In *Proceedings of IEEE Winter Conference on Applications of Computer Vision*, pages 1137–1144, 2015.

[18] S. Pellegrini, K. Schindler, and D. Nardi. A generalization of the ICP algorithm for articulated bodies. In *Proceedings of British Machine Vision Conference*, pages 87.1–87.10, 2008.

[19] O. Sorkine, D. Cohen-Or, Y. Lipman, M. Alexa, C. Rössl, and H.-P. Seidel. Laplacian surface editing. In *Proceedings of Eurographics/ACM SIGGRAPH Symposium on Geometry Processing*, pages 175–184, 2004.

[20] J. Tang, L. Shao, and X. Zhen. Robust point pattern matching based on spectral context. *Pattern Recognition*, 47(3):1469–1484, 2014.

[21] A. Weiss, D. Hirshberg, and M. J. Black. Home 3D body scans from noisy image and range data. In *Proceedings of IEEE International Conference on Computer Vision*, pages 1951–1958, 2011.

[22] J. Yang, J. P. Williams, Y. Sun, R. S. Blum, and C. Xu. A robust hybrid method for nonrigid image registration. *Pattern Recognition*, 44(4):764–776, 2011.

[23] M. Ye, X. Wang, R. Yang, R. Liu, and M. Pollefeys. Accurate 3D pose estimation from a single depth image. In *Proceedings of IEEE International Conference on Computer Vision*, pages 731–738, 2011.

[24] M. Ye and R. Yang. Real-time simultaneous pose and shape estimation for articulated objects using a single depth camera. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 2353–2360, 2014.

[25] A. L. Yuille and N. M. Grzywacz. A mathematical analysis of the motion coherence theory. *International Journal of Computer Vision*, 3(2):155–175, 1989.

[26] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.

[27] Y. Zheng and D. Doermann. Robust point matching for nonrigid shapes by preserving local neighborhood structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):643–649, 2006.