統計的形状モデルを利用した大規模
点群レジストレーションとその高速化

広瀬 修†

金沢大学理工研究域†

1. はじめに

点群位置合わせ問題は，物体の形状を表現する２つの点集合間に対し，１つの点群をもう一方の点群に移す写像を求める問題である．点群位置合わせ問題は想定する写像の種類に応じて剛体変換と非剛体変換のものに分類され，最近では非剛体変換に基づいた点群位置合わせが，その柔軟さのため非常に活発に研究されている．

Coherent point drift (CPD) は非剛体変換に基づいた点群レジストレーション手法の代表的な手法である [Myronenko 2010]．CPD の成功の要因としてまず挙げられるのが，外れ値への耐性である．ここで，外れ値とは点群によって表現される形状とは無関係に存在する点とする．CPD は点群位置合わせ問題を混合確率分布の推定問題として定義する．その際，混合分布の構成分布の１つとして外れ値の分布を明示的に与えることが，外れ値への耐性の主要な要因である．CPD のもう１つの成功要因として挙げられるのが，非剛体換される点群に対する「変位場の滑らかさ」である．変位場の滑らかさとは，非剛体変換される点群を構成する任意の点の変位と，その他の点の変位が，その距離が近ければ近いほど相関するとした仮定である．この仮定は非常に自然な仮定であるため，CPD は多くの点群位置合わせ問題において精度の高い位置合わせ結果を与える．一方で，変位場の滑らかさの仮定が適切ではない場合，CPD は容易に位置合わせに失敗する．例えば，人間の手の形状マッチングを行う場合，人差し指と中指を構成する点は比較的近くに位置するが，その動きは逆相関する傾向があり，このような場合には変位場の滑らかさの仮定だけでは不十分であるためである．

この問題を克服する方法の１つとして挙げられるのが，教師あり学習に基づく方法である．もし人差し指と中指の動きが逆相関する傾向にあることを事前に知っていれば，その知識を位置合わせアルゴリズムに組み込むことにより，高精度の位置合わせが期待できるからである．今回，新たに開発した教師あり学習法に基づいた点群位置合わせ手法に対する位置合わせ性能の評価について報告する．開発手法は外れ値の分布を構成分布の１つとして持つ混合分布に基づいているため，外れ値への優れた耐性を有する．また，物体の形状変化モデルに，訓練データから得られる事前知識を組み込むため，変位場の滑らかさの問題を自然に解決することができる．講演では大規模な点群データに対応するための高速化についても議論を行う．

2. 実験

開発した位置合わせ手法と代表的な点群位置合わせ手法である CPD と Thin Plate Spline Robust Point Matching (TPS-RPM) [Chui 03] の性能の比較を行った．使用したデータは IMM hand データである[Stegmann 2002]．このデータは４０種類の人間の手の画像に対し，手と背景の境界部分に５６個の特徴点を人手で打点したものである．各々の特徴点は４０種類の画像で対応関係がとられるように打点されている．図１が IMM hand データ中の点群番号６に対する CPD，TPS-RPM および開発手法の位置合わせ結果を表す．図１の１段目が推定されるべき正解の手の形状を表し，２段目が位置合わせの対象となるデータで，手法の頑健性を検証するため正解データに４種類の改変を施したものである．左から順に (1)点の複製，(2)欠損，(3)外れ値の付加，(4)回転の改変を表す．図１の３段目の赤色で示された点が最適化の初期形状，すなわち，平均形状を表す．３～５段目がそれぞれ提案手法，CPD，TPS-RPM の適用結果を表す．データ (1)，(2)，(3)，(4)の全てで提案手法が正解とほぼ同一の形状の推定に成功した．CPD と TPS-RPM については，全てのデータで親指以外の指が細くなる現象が見られた．これは，変位場の滑らかさの欠点の１つを表している．ある指を形成する点が，近隣に存在する他の指を形成する点ともその変位が相関することが原因である．データ(1) に対しては全ての手法で概ね正解の形状を空いてしたが，CPD と TPS-RPM には指が細くなる現象が見られた．データ(2) に対して提案手法は欠損領域の点群の推定に成功し

た．一方で CPD と TPS-RPM は親指を含む全ての　指が短く描画された．

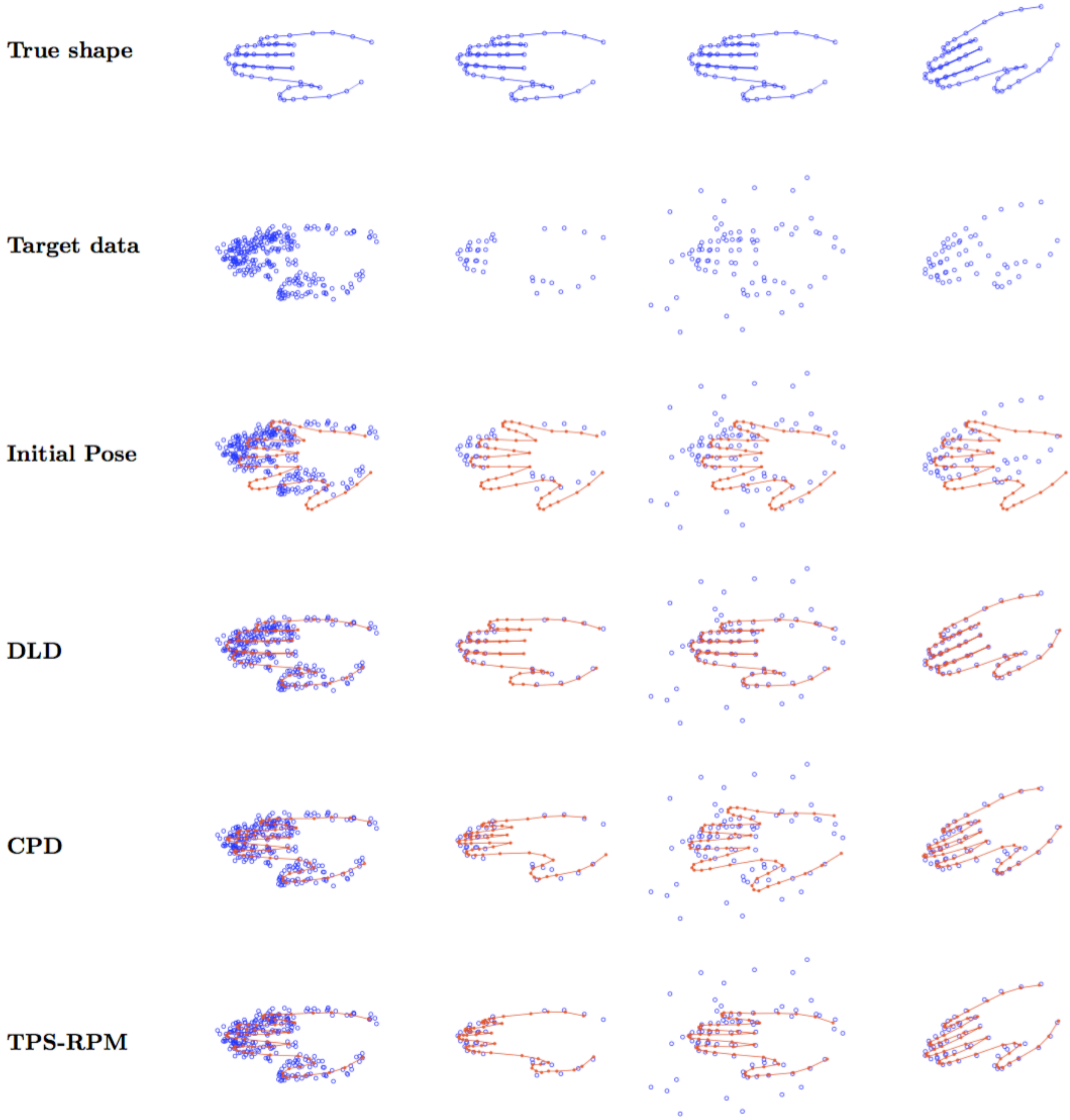

図 1．提案手法，CPD，TPS-RPM の比較．１段目の図は推定されるべき正解の形状を表す．２段目の図は位置合わせの対象となるデータ．３段目の赤の点群が最適化の初期形状を表す．手法の頑健性を検証するため正解データに４種類の改変を施した．左から順に (1)点の複製，(2)欠損，(3)外れ値の付加，(4)回転の改変を表す．３～５段目がそれぞれ提案手法，CPD，TPS-RPM の適用結果を表す．

Robust point set registration using a statistical shape model

Osamu Hirose

Institute of Science and Engineering, Kanazawa University,
Kakuma, Kanazawa, Ishikawa 920-1192, JAPAN.

Abstract. Point set registration is a problem of finding point-by-point correspondences between two point sets, each of which characterizes an object shape. Many of the state-of-the-art algorithms to solve the problem are based on the assumption of the smooth displacement field, which enforces neighbor points to move coherently. The assumption is reasonable in many situations and the algorithms often solve point set registration problems elegantly. However, these algorithms often fail if the assumption is inappropriate. In the case of registering two hand poses, for example, moves of the index finger and the middle finger tend to be negatively correlated. A key to overcoming the issue of the smooth displacement field is the use of prior knowledge of object geometry. If we know in advance the fact that the index and middle fingers tend to move with negative correlation, the issue can be avoided by incorporating the information into the registration algorithm to be designed. In this paper, we propose a novel point set registration algorithm based on a statistical shape model, a supervised learning technique for learning shape variations of an object. An effectiveness of the algorithm is presented through comparisons with the state-of-the-art point set registration algorithms using the data of human hand poses with various types of artifacts.

1 Introduction

Point matching is a problem of finding point-by-point correspondence between two point sets where each of point sets characterizes a geometry of an object. Finding such geometrical correspondence of point sets are being actively studied in the field of the image recognition and the computer vision. One major class of the point matching problems is point set registration, the problem of finding a transformation between two point sets including point-by-point correspondence. The point set registration problems are roughly classified into two classes according to transformation models: rigid and non-rigid transformations. The rigid transformation model is defined as a linear map which preserves relative positions of points in a floating point set, i.e. scaling, rotation, and translation. The rigid point set registration problem is a relatively simple problem and has been intensively studied [2, 1, 23, 7, 18]. The non-rigid registration is a more complex problem that transforms a shape of an object geometry. Typical transformation models used for the point set registration problems are the thin-plate spline [5, 6, 14, 3, 27] and the radial basis function [21, 20, 16, 17]. These methods are differently classified according to definitions of the point set registration problems: energy minimization [5, 6, 27], and probabilistic density estimation with Gaussian mixture model [21, 20, 14, 3, 16, 17].

One key to the success of these registration methods is a robustness for outliers, points irrelevant to the true object geometry. There are several approaches to deal with outliers, statistical analysis for distances of correspondent points [28, 11, 25], soft assignments [22, 6], trimming point sets through iterative random sampling [4], kernel correlation [26], explicit probabilistic modeling of outliers [21, 20, 16, 17], the use of a robust estimator: the L_2E estimator [14] and a scaled Geman-McClure estimator [29]. The second key to the success of the registration methods is an assumption of the smooth displacement field, which enforces neighbor points to move coherently. The smoothness of the displacement field is imposed by a regularization technique defined as a penalty term for energy minimization [5, 6, 27] and for log-likelihood functions [21, 20, 14, 3, 16, 17]. Owing to the assumption of the smooth displacement field, such non-rigid registration algorithms find transformations with sufficient global flexibility while local topology of a point set is preserved. The smoothness of the displacement field, however, becomes a significant drawback if the assumption of the true displacement field is not smooth. In the case of matching two hand shapes, for example, tips of the index finger and the middle finger are located closely whereas they usually move with a negative correlation.

To address the issue of the smooth displacement field, a promising direction is the use of prior knowledge of object geometry. If we know the fact that the index and middle fingers tend to move with a negative correlation in advance, the issue is expected to be relaxed by incorporating the information into registration algorithms to be designed. One approach to incorporating such prior information is the use of a kinematic motion for articulated objects such as human body [19, 13, 9, 8]. These methods show promising results, but they cannot be applied to objects with no kinematics. The second approach to overcoming the issue of the smooth displacement field is the use of partial correspondence across two point sets [15, 10]. These methods also show promising results, but the better performance is not expected if the partial correspondence is not available. The third approach to addressing the issue of the smooth displacement field. One candidate of such supervised learning techniques is a statistical shape model [24, 12] which describes the mean shape and statistical variation of geometrical objects. Shape variations represented by statistical shape models are constructed from shape statistics of landmark displacements which are usually obtained manually. Therefore, moves of neighbor landmarks are not assumed to be correlated, and moves of distant landmarks are allowed to be dependent, unlike the smooth displacement field. Also, statistical shape models do not require any physical models such as object kinematics.

Contribution of this work

In this article, we propose a novel non-rigid point set registration algorithm, which we call dependent landmark drift. The problem we aim at solving is point set registration problems under the condition that (i) complete point-by-point correspondence across multiple point sets are available as a set of training data, and (ii) no partial correspondence across two point sets to be registered is available. The proposed algorithm is based on a Gaussian mixture model with robustness for outliers and a statistical shape model, a supervised learning technique for learning shape variations of an object. The statistical shape model is constructed from a set of training data and does not require the assumption of the smooth displacement field. Therefore, the difficulty in the registration of non-coherent neighbor points is addressed without losing robustness for outliers. The proposed algorithm is specialized for registration problems between the mean shape and a point set, and works much more efficiently especially if the assumption of the smooth displacement field is inappropriate. An effectiveness of the proposed algorithm will be presented through the comparison with the state-of-the-art point set registration algorithms.

2 Methods

In this section, we propose a novel point set registration algorithm based on a Gaussian mixture model with a statistical shape model. we first describe a definition of a statistical shape model, and then introduce a GMM-based point set registration technique [20]. Finally, a novel algorithm which addresses the issue of non-coherent neighbor points is presented.

2.1 Statistical shape model

we begin with definitions of a landmark and a shape required to define a statistical shape model. To obtain shape statistics from multiple geometric objects, it is essential to define correspondent points across the geometric objects. These points of correspondence are called *landmarks*. A *shape* is typically defined as a set of landmarks for one of the geometric objects with scale, rotation and translation effects removed [24]. Unlike this definition, we define a shape simply as a set of landmarks, that is, scale, rotation, and translation effects are not essentially removed for the purpose of point set registration. Note here that a point set and a shape are distinguished from each other in that (1) shapes are composed of the same number of points whereas the number of points in a point set is generally different from other point sets, and (2) points in shapes are correspondent across all the shapes whereas points in multiple point sets are not correspondent. Statistical shape models, are constructed from training shapes, i.e. multiple point sets with point-by-point correspondence.

Definition of a statistical shape model

Statistical shape model (SSM) is an expression of a geometrical shape and its statistical variations for an object. Since definitions of SSMs diverge according to the aim of applications or the method of construction [12], we introduce a definition based on the principal component analysis (PCA) [24] which is sufficient to describe the algorithm to be proposed. Suppose a shape is composed of M landmarks $(v_1, \cdots, v_M)$, each of which lie in a

D-dimensional space. Then, the shape is represented as a vector $\mathbf{v} = (v_1^T, \cdots, v_M^T)^T \in \mathbb{R}^{MD}$. The PCA-based statistical shape model is defined as follows

$$\mathbf{v} = \mathbf{u} + \sum_{k=1}^{K} z_k \mathbf{h}_k + \mathbf{w}, \tag{1}$$

where $\mathbf{u} = (u_1^T, \cdots, u_M^T)^T \in \mathbb{R}^{MD}$ is the mean shape, $\mathbf{h}_k \in \mathbb{R}^{MD}$ is the kth leading shape variation, $z_k \in \mathbb{R}$ is the kth weight corresponding to the kth shape variation, K is the number of shape variations, and $\mathbf{w} \in \mathbb{R}^{MD}$ is a residual vector. To separate shape variations to a maximum extent, shape variations $\mathbf{h}_1, \cdots, \mathbf{h}_M$ are assumed to satisfy the orthonormality condition $\mathbf{h}_i^T \mathbf{h}_j = \delta_{ij}$ where δ_{ij} is Kronecker's delta.

Estimation of shape parameters

The shape parameters $\mathbf{u}$, $\mathbf{h}_k$, and, z_k are unknown and should be estimated from multiple shapes $\mathbf{v}_1, \cdots, \mathbf{v}_B$. The mean shape $\mathbf{u}$ is simply estimated as the average of sample shapes. Suppose $\mathbf{C} \in \mathbb{R}^{MD \times MD}$ is a shape covariance matrix defined as

$$\mathbf{C} = \frac{1}{B-1} \sum_{j=1}^{B} (\mathbf{v}_j - \bar{\mathbf{v}})(\mathbf{v}_j - \bar{\mathbf{v}})^T, \tag{2}$$

where $\bar{\mathbf{v}}$ is the sample average of the shapes $\mathbf{v}_1, \cdots, \mathbf{v}_B$. The covariance matrix $\mathbf{C}$ represents statistical dependencies for landmark displacements. The kth shape variation $\mathbf{h}_k$ and the corresponding weight z_k in the equation (1) can be estimated as the kth eigenvector and the kth eigenvalue of $\mathbf{C}$, respectively.

2.2 Gaussian mixture model for point set registration

we summarize a Gaussian mixture modeling approach for solving point set registration problems proposed by Myronenko et al. [20] since this approach is the basis of the algorithm to be proposed. They defined a registration problem of two point sets as a problem of probabilistic density estimation where one point set is composed of centroids for a Gaussian mixture model (GMM) and the other point set is samples generated from the GMM. Suppose $x_n \in \mathbb{R}^D$ and $y_l \in \mathbb{R}^D$ are the nth element in a target point set $X = \{x_1, \cdots, x_N\}$ and the lth element in a floating point set $Y = \{y_1, \cdots, y_L\}$, respectively. Their definition of the mixture model for the alignment of the two point sets X and Y is

$$p(x_n; \Theta) = \omega \cdot p_{\text{outlier}}(x_n) + (1 - \omega) \cdot \sum_{l=1}^{L} p(l) p(x_n | l; \Theta), \tag{3}$$

where Θ is a set of parameters of the mixture model, ω is a Bernoulli probability for outliers, and p_{outlier} is a distribution of outliers. Prior distributions of inliers $p(l)$ and outliers $p_{\text{outlier}}(x_n)$ are defined as uniform distributions $p(l) = 1/L$ and $p(x_n) = 1/N$, respectively. The inlier distribution $p(x_n | l; \Theta)$ is defined as a Gaussian distribution

$$p(x_n | l; \Theta) = \frac{1}{(2\pi\sigma^2)^{D/2}} \exp\Big(- \frac{||x_n - \mathcal{T}(y_l; \theta)||^2}{2\sigma^2} \Big), \tag{4}$$

where σ^2 is the variance of the Gaussian distribution, $\mathcal{T}(y_l; \theta)$ is a transformation model for the floating point y_l with a set of parameters θ, and $\Theta = (\theta, \sigma^2)$ is a set of parameters of the GMM. The set of parameters Θ can be estimated by the maximum likelihood estimation [20]. Since the analytic solution of the maximum likelihood estimation for the GMM, defined as the equations (3) and (4), is not available, the EM algorithm is used for searching a local maximum of the likelihood function. The EM algorithm iteratively improves a solution by updating a lower bound of a log-likelihood function, called Q-function. Given a parameter set $\bar{\Theta} = (\bar{\sigma}^2, \bar{\theta})$, the Q-function of the GMM is derived as

$$Q(\Theta, \bar{\Theta}) = \frac{N_P D}{2} \log \sigma^2 + \frac{1}{2\sigma^2} \sum_{n=1}^{N} \sum_{l=1}^{L} p(l | x_n, \bar{\Theta}) ||x_n - \mathcal{T}(y_l; \theta)||^2. \tag{5}$$

where $N_P = \sum_{n=1}^{N} \sum_{l=1}^{L} p(l|x_n, \bar{\Theta}) \leq N$ is the effective number of matching points and $p(l|x_n, \bar{\Theta})$ denotes a posterior probability of the mixture component. This posterior probability can be calculated as

$$p(l|x_n; \bar{\Theta}) = \frac{(1-\omega)p(x_n|l; \bar{\Theta})}{\omega \frac{1}{N} + (1-\omega)\frac{1}{L} \sum_{l'=1}^{L} p(x_n|l'; \bar{\Theta})}. \tag{6}$$

Therefore, a solution of the point set registration problem is obtained by iterating the following procedure: (i) updating the posterior probability $p(l|x_n; \bar{\Theta})$, (ii) finding $\hat{\Theta}$ which maximizes the Q-function for Θ given the current parameter set $\bar{\Theta}$, and (iii) replacing the current parameter set $\bar{\Theta}$ with the maximizer $\hat{\Theta}$ of the Q-function. This procedure is iterated until a suitable convergence criterion is satisfied.

2.3 Dependent landmark drift

We here propose a novel registration algorithm specialized for registering the mean shape and a point set, which we call dependent landmark drift (DLD). The algorithm is based on the same GMM framework as the CPD algorithm. A main difference between CPD and DLD is the difinition of transformation models for a floating point set: CPD uses a motion coherence while DLD uses a statistical shape model. We first describe that statistical shape models can be utilized as a transformation model of the GMM-based point set registration. Suppose $h_{mk} \in \mathbb{R}^D$ is the subvector of the kth shape variation $\mathbf{h}_k$, corresponding to the mth landmark $u_m \in \mathbb{R}^D$ in the mean shape $\mathbf{u} \in \mathbb{R}^{MD}$. We also denote K shape variation vectors corresponding to the mth landmark by a D-by-K matrix $H_m = (h_{m1}, \cdots, h_{mK}) \in \mathbb{R}^{D \times K}$. Then, the statistical shape model (1) is denoted by a point-by-point transformation model suitable for point set registration problems:

$$\begin{aligned} v_m &= \mathcal{T}_{H_m}(u_m; z) + w_m \\ &= u_m + H_m z + w_m, \end{aligned} \tag{7}$$

where $z = (z_1, \cdots, z_K) \in \mathbb{R}^K$ is a weight vector for K shape variations, and $w_m \in \mathbb{R}^D$ is a subvector of the residual vector $\mathbf{w}$ corresponding to the landmark u_m. A merit of using statistical shape models for point set registration problems is that moves of landmarks are estimated based on *statistical dependency* of landmark displacements without the assumption of the smooth displacement field. That is, neighbor landmarks are not enforced to be coherent and moves of distant landmarks are allowed to be dependent. Therefore, we call the algorithm *dependent landmark drift*. We note here that the affine transformation $\mathcal{T}_{H_m}$ is much more flexible than that of the affine CPD since all of the transformation matrices $H_1, \cdots$, and H_M are identical for the affine CPD and are not identical for the statistical shape models.

We then derive the DLD algorithm for solving a point set registration problem for a point set $X = \{x_1, \cdots, x_N\}$ and the mean shape $U = \{u_1, \cdots, u_M\}$, also denoted by single vector notation $\mathbf{u} = (u_1^T, \cdots, u_M^T)^T \in \mathbb{R}^{MD}$. Based on the GMM framework, we set the mean shape U as the floating points and set a statistical shape model $\mathcal{T}_{H_m}(u_m; z)$ trained in advance as a transformation model. Then, we define the Q-function of the point set registration problem as

$$Q(\Theta, \bar{\Theta}) = \frac{N_P D}{2} \log \sigma^2 + \frac{1}{2\sigma^2} \sum_{n=1}^{N} \sum_{m=1}^{M} p_{mn} ||x_n - (u_m + H_m z)||^2 + \gamma ||z||^2, \tag{8}$$

where $\Theta = (z, \sigma^2)$, $p_{mn} = p(m|x_n; \Theta)$, and $N_P = \sum_{n=1}^{N} \sum_{m=1}^{M} p_{mn}$. A regularization term was added for z in order to avoid searching extreme shapes where $\gamma > 0$ is a parameter which controls the search space of z, meaning that the resulting transformed shape becomes closer to the mean shape as γ increases. Since the analytic solution of the simultaneous maximization of the Q-function for z and σ^2 is not available, we optimize z and σ^2 separately, likely to the CPD algorithm. By optimizing z given σ^2, or optimizing σ^2 given z, we have

$$\hat{z} = \Big\{ \sum_{n=1}^{N} \sum_{m=1}^{M} p_{mn} (H_m^T H_m + \gamma I) \Big\}^{-1} \Big\{ \sum_{n=1}^{N} \sum_{m=1}^{M} p_{mn} H_m^T (x_n - u_m) \Big\} \tag{9}$$

$$\hat{\sigma^2} = \frac{1}{N_P D} \sum_{n=1}^{N} \sum_{m=1}^{M} p_{mn} ||x_n - (u_m + H_m z)||^2. \tag{10}$$

Therefore, a solution of the point set registration problem is obtained by iterating the following procedure: (i) updating the posterior probability $p(m|x_n; \bar{\Theta})$, (ii) finding $\hat{\Theta}$ which maximizes Q-function for Θ given the current parameter set $\bar{\Theta}$, and (iii) replacing the current parameter set $\bar{\Theta}$ with the maximizer $\hat{\Theta}$ of the Q-function.

3 Experiments

In this section, we evaluate registration performance of the proposed algorithm through comparisons with the state-of-the-art point set registration algorithms, CPD [20] and TPS-RPM [5]. We used the IMM hand data [25]. This data includes 40 shapes, and each of the shapes was obtained from a 2D image of a human hand by placing 56 landmarks manually. That is, all of the points included in the 40 shapes are correspondent across all the shapes. As a pre-alignment, we translated each of the 40 hand poses as follows: the average point in a hand pose becomes the origin in the two-dimensional coordinate space.

To evaluate registration performances of DLD more precisely, we compared DLD, CPD, and TPS-RPM using the same IMM hand data under various conditions. As the true target poses to be estimated, we used three hand poses, No.6, No.9 and No.38 which are clearly different poses from the mean hand, shown in the top row of Figure 1. As a set of training data for DLD, we used 39 hand poses with each target hand pose removed from all the 40 hand poses. As in the above demonstration, we generated the four types of target data for the three hand poses: (a) 20-times replication of target points with dispersion, (b) random deletion of target points, (c) addition of outliers which follows uniform distributions, and (d) rotation of the whole shape. In generating the four types of target data, we changed parameters of the data generation: (a) standard deviation of the Gaussian distribution for replicating target points, (b) missing rate, (c) signal-to-noise ratio, and (d) rotation angle. To repeat experiments and reduce an influence of the randomness on the registration performance, target point sets were generated 20 times for (a), (b) and (c) in the same setting. For (d), a single set of target points was generated for each rotation angle since the randomness does not exist for the rotation. DLD, CPD, and TPS-RPM were applied to each target data, and a registration accuracy was calculated for each target data. For the target data (a), (b), and (c), averages and standard errors of registration accuracy were calculated. The registration accuracy was defined as the rate of the correct matching between a true target pose and a deformed mean hand after registration. Here, the point-by-point correspondence was estimated by using the nearest-neighbor matching. We used the default parameters for CPD and TPS-RPM, and we fixed $\omega = 0.9$ and $K = 10$ and changed the regularization parameter γ to $10^{-3}, 10^{-4}$, and 10^{-5} for DLD.

The second row in Figure 1 shows the results of the comparison for the target hand poses with replicated target points. The x-axis represents standard deviation for replicating target points while the y-axis registration accuracy. For all the three target poses, The registration performance of DLD was insensitive to the regularization parameter γ. For hand No.6, DLD and CPD worked comparatively whereas DLD outperformed CPD and TPS-RPM for hand No.9 and hand No.38 for almost all cases.

The results for target poses with random missing points are shown in the third row of Figure 1. The x-axis represents the missing rate, and the y-axis represents the registration accuracy. DLD with $\gamma = 10^{-4}$ worked most stable for all the hand poses while DLD with $\gamma = 10^{-3}, 10^{-5}$ was less accurate than CPD in some cases, especially for the high missing rate. The registration accuracy of DLD with $\gamma = 10^{-5}$ rapidly decreased for all of the target poses in comparison with DLD with $\gamma = 10^{-3}, 10^{-4}$ as the missing rate became large. This rapid decrease in registration accuracy suggests that shape prior information is insufficient to impute missing regions. In other words, missing region can be reasonably imputed by shape prior information if the regularization parameter γ was appropriately chosen. DLD with $\gamma = 10^{-3}$ was less accurate than DLD with $\gamma = 10^{-4}, 10^{-5}$ for low missing rate, suggesting that large γ sometimes becomes an obstacle to accurate registration.

The fourth row in Figure 1 shows the results for target poses with outliers which follows a two-dimensional uniform distribution with $x_1 \in [-0.6, 0.6]$ and $x_2 \in [-0.6, 0.6]$. The x-axis of each figure represents signal-to-noise ratio, and the y-axis represents registration accuracy. DLD achieved the best registration performance in many cases, showing a robustness of DLD for outliers. As with cases of random missing points, an importance of the regularization parameter γ can be observed. DLD with $\gamma = 10^{-3}$ tends to be less accurate than DLD with $\gamma = 10^{-4}, 10^{-5}$ for low signal-to-noise ratio while the decrease in registration performance of DLD with $\gamma = 10^{-3}$ was relatively small for high signal-to-noise ratio.

The fifth row in Figure 1 shows the results for rotated target poses. The x-axis and the y-axis represent rotation angle and registration accuracy, respectively. For these data, registration accuracy of TPS-RPM were considerably stable and were not affected by rotation angles at least in range $[-\pi/3, \pi/3]$ for all the three target poses. For the rotation, DLD was less robust than TPS-RPM and was comparative with CPD, but achieved the best registration performance for relatively low rotation angles.

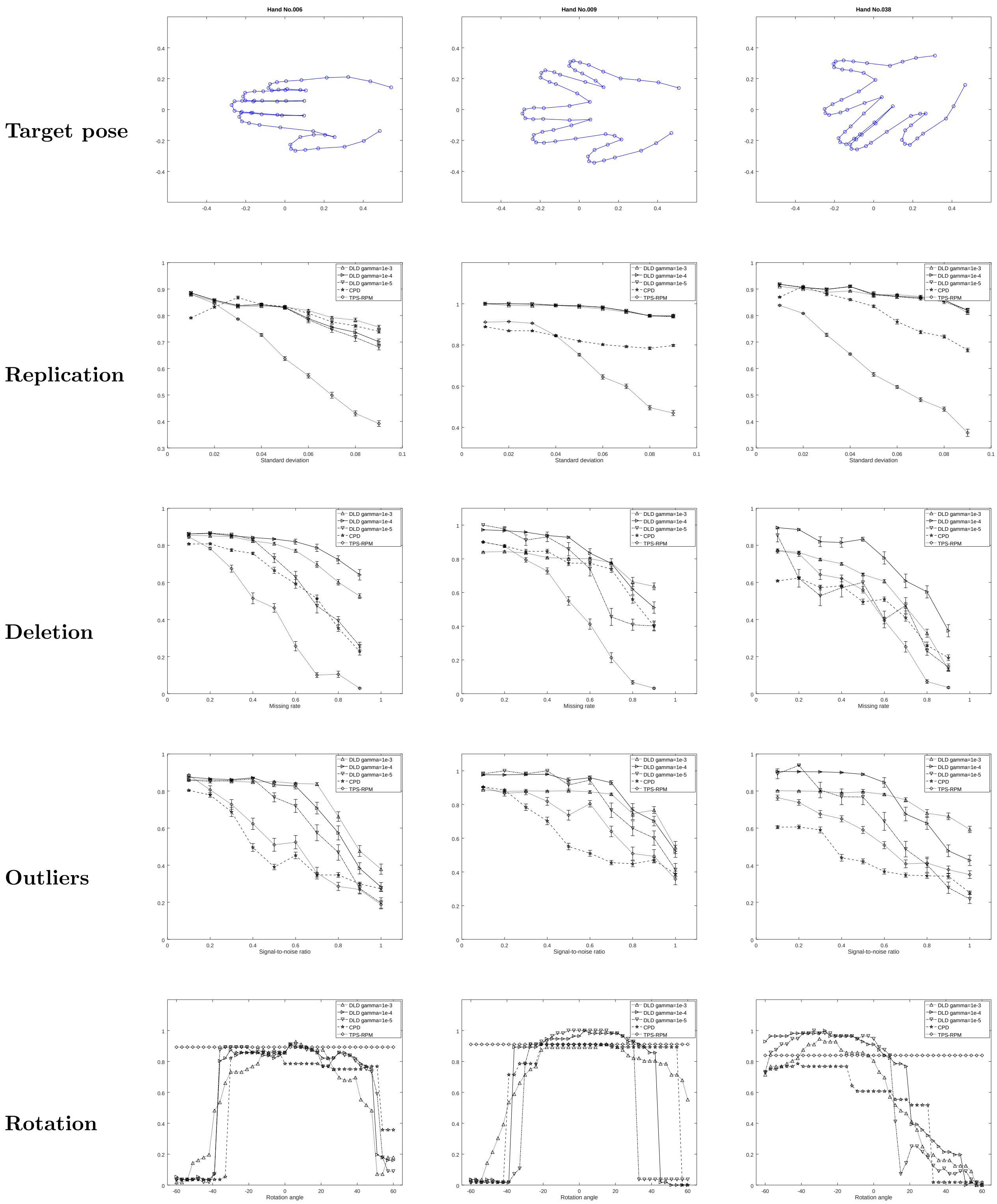

Figure 1: Comparisons of registration performance for DLD, CPD, and TPS-RPM.

4 Conclusion

Many of the state-of-the-art point set registration algorithms are based on the assumption of the smooth displacement field, meaning that moves of neighbor points in a floating point set are correlated, and local geometry of the point set is preserved. Owing to the assumption, registration problems are elegantly solved by these algorithms in many cases. However, these algorithms often fail to register two point sets if the assumption is inappropriate. For example, in a case of registering two hand poses, points in the index finger and the middle finger tend not to be correlated.

A key to overcoming this issue is the use of supervised learning techniques. If we know in advance the fact that moves of the index finger and the middle finger are negatively correlated, we do not have to rely on the smooth displacement field. Based on this idea, we proposed a novel point set registration algorithm using a statistical shape model. The algorithm is specialized for registering the mean shape and a point set, and works efficiently even if the assumption of the smooth displacement field is not expected.

To evaluate registration performance of the proposed algorithm, we compared the algorithm with CPD and TPS-RPM using the IMM hand data with four types of modifications: replication of target points with dispersion, random deletion of target points, addition of outliers, and rotation of the whole shape. The proposed algorithm outperformed CPD and TPS-RPM in many cases for the data, showing a robustness of the proposed algorithm for various types of point set registration problems.

References

[1] Paul Besl and Neil McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[2] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.

[3] Jun Chen, Jiayi Ma, Changcai Yang, Li Ma, and Sheng Zheng. Non-rigid point set registration via coherent spatial mapping. *Signal Processing*, 106:62–72, 2015.

[4] Dmitry Chetverikov, Dmitry Stepanov, and Pavel Krsek. Robust Euclidean alignment of 3D point sets: The trimmed iterative closest point algorithm. *Image and Vision Computing*, 23(3):299–309, 2005.

[5] Haili Chui and Anand Rangarajan. A feature registration framework using mixture models. In *Proceedings of IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA'00)*, pages 190–197. IEEE Comput. Soc, 2000.

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

[7] Andrew W Fitzgibbon. Robust registration of 2D and 3D point sets. *Image and Vision Computing*, 21(13-14):1145–1153, dec 2003.

[8] Song Ge and Guoliang Fan. Non-rigid Articulated Point Set Registration for Human Pose Estimation. In *Proceedings of IEEE Winter Conference on Applications of Computer Vision 2015 (WACV'15)*, pages 94–101. IEEE Comput. Soc, jan 2015.

[9] Song Ge and Guoliang Fan. Non-rigid articulated point set registration with Local Structure Preservation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops 2015 (CVPR'15)*, pages 126–133. IEEE Comput. Soc, jun 2015.

[10] Vladislav Golyanik, Bertram Taetz, Gerd Reis, and Didier Stricker. Extended coherent point drift algorithm with correspondence priors and optimal subsampling. In *Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV'16)*, pages 1–9. IEEE, mar 2016.

[11] Sébastien Granger and Xavier Pennec. Multi-scale EM-ICP: A Fast and Robust Approach for Surface Registration. In *Proceedings of the 4th European Conference on Computer Vision (ECCV'02)*, pages 418–432. Springer-Verlag, 2002.

[12] Tobias Heimann and Hans Peter Meinzer. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*, 13(4):543–563, 2009.

[13] Radu Horaud, Florence Forbes, Manuel Yguel, Guillaume Dewaele, and Jian Zhang. Rigid and articulated point registration with expectation conditional maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3):587–602, 2011.

[14] Bing Jian and Baba C. Vemuri. Robust point set registration using Gaussian mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1633–1645, 2011.

[15] Ivan Kolesov, Jehoon Lee, Gregory Sharp, Patricio Vela, and Allen Tannenbaum. A Stochastic Approach to Diffeomorphic Point Set Registration with Landmark Constraints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):238–251, 2016.

[16] Jiayi Ma, Weichao Qiu, Ji Zhao, Yong Ma, Alan L. Yuille, and Zhuowen Tu. Robust L2E estimation of transformation for non-rigid registration. *IEEE Transactions on Signal Processing*, 63(5):1115–1129, 2015.

[17] Jiayi Ma, Ji Zhao, and Alan L Yuille. Non-Rigid Point Set Registration by Preserving Global and Local Structures. *IEEE Transactions on Image Processing*, 25(1):53–64, 2016.

[18] A Makadia, A.I. Patterson, and K Daniilidis. Fully Automatic Registration of 3D Point Clouds. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'06)*, pages 1297–1304. IEEE Comput. Soc, 2006.

[19] Diana Mateus, Radu Horaud, David Knossow, Fabio Cuzzolin, and Edmond Boyer. Articulated shape matching using Laplacian eigenfunctions and unsupervised point registration. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition 2008 (CVPR'08)*, pages 1–8. IEEE Comput. Soc, jun 2008.

[20] Andriy Myronenko and Xubo Song. Point set registration: Coherent point drifts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(12):2262–2275, 2010.

[21] Andriy Myronenko, Xubo Song, and Miguel Á Carreira-Perpiñán. Non-rigid point set registration: Coherent Point Drift. In *Proceedings of Advances in Neural Information Processing Systems 19 (NIPS'06)*, pages 1009–1016. MIT press, 2006.

[22] Anand Rangarajan, Haili Chui I, and Fred L Bookstein. The Softassign Procrustes Matching Algorithm. *Lecture Notes in Computer Science*, 1230:29–42, 1997.

[23] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pages 145–152. IEEE Comput. Soc, 2001.

[24] Mikkel B. Stegmann and David Delgado Gomez. A brief introduction to statistical shape analysis, 2002.

[25] C Stewart, L Tsai, and B Roysam. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):1379–1394, 2003.

[26] Yanghai Tsin and Takeo Kanade. A correlation-based approach to robust point set registration. In *Proceedings of the 6th European Conference on Computer Vision (ECCV'04)*, pages 558–569. Springer-Verlag, 2004.

[27] Yang Yang, Sim Heng Ong, and Kelvin Weng Chiong Foong. A robust global and local mixture distance based non-rigid point set registration. *Pattern Recognition*, 48(1):156–173, 2015.

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

[29] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast Global Registration. In *Proceedings of the 14th European Conference on Computer Vision (ECCV'16)*. Springer-Verlag, 2016.