A new joint clustering and diffeomorphism estimation algorithm for non-rigid shape matching

I M 1 Institution 1
I M 2 Institution2
I M 3 Institution 3

Abstract

Matching shapes parameterized as unlabeled point-sets is a challenging problem since we have to solve for point correspondences in a non-rigid setting. Previous work on this problem such as modal matching, linear assignment, shape contexts etc. has focused more on the correspondence aspect and not on the non-rigid deformations. The principal motivation for the present work is to establish a distance measure between shapes on a shape manifold. A prerequisite for achieving this goal is the diffeomorphic *matching of point-sets. We show that a joint clustering and diffeomorphism estimation strategy is capable of simultaneously estimating correspondences and a diffeomorphism between unlabeled point-sets. Cluster centers for the two point-sets having the same label are always in correspondence. Essentially, as the cluster centers evolve during the iterations of an incremental EM algorithm, we estimate a diffeomorphism between the two sets of cluster centers. We apply our algorithm to 2D corpus callosum shapes.*

1 Introduction

In recent years, we have seen considerable interest in the application of statistical shape analysis to problems in medical image analysis, computer graphics and computer vision. Regardless of whether shapes are parameterized by point-sets, curves, graphs etc., the fundamental problem of estimating shape correspondence in a non-rigid setting remains. We are particularly interested in the unlabeled point-set parameterization since statistical shape analysis of point-sets is very mature [1]. Means, covariances and even probability distributions on shape manifolds can be defined and estimated. However, the correspondence problem is particularly acute for the point-set representation. This is mainly due to the fact that the topology information, or concept of neighborhood which narrows down the set of allowed correspondences is missing in the case of discrete point-sets while obviously available (and useful) for example, in the curve parameterization.

The considerable amount of previous work on the estimation of correspondences and deformations between unlabeled point-sets is hard to summarize. Rather than attempt such a foolhardy task and incur the wrath of outraged reviewers, we instead focus on the fact that there is very little previous work on estimating correspondences and *diffeomorphisms* between unlabeled point-sets. Our motivation for the present work is driven by two issues: i) diffeomorphisms are fundamental in shape manifold theory and ii) our ultimate goal is to build probability distributions on shape manifolds for shapes parameterized as unlabeled point-sets. When deformations are represented using splines (B-splines, thin-plate splines and the like), there is no guarantee that the mapping is a bijection. Hence the non-rigid mapping may have local folds and this causes point features to be incorrectly mapped.

In parallel to the above work on point correspondences is the study of diffeomorphic landmark matching in [2] and [3]. In both the above works, a flow field is used to generate diffeomorphisms that maps source landmarks to target landmarks. In [2], the flow and the diffeomorphism are constrained by the transport equation. The deformation energy is expressed in term of the smoothness of the flow field.

$$E(v) = \int_0^1 \int_\Omega \parallel Lv(x,t) \parallel^2 dxdt \qquad (1)$$

where $v(x,t)$ is a velocity field (as opposed to a displacement field) and t is a time parameter. The smoothness operator L is standard and we use a thin-plate spline [4] in all our experiments.

They also studied the inexact matching problem between two sets of landmarks $x_n,\ y_n, \forall n \in \{1,\ldots,N\}$. When the images of the source landmarks are allowed to not precisely match target landmarks, a second term which evaluates how well the two sets of landmarks match is introduced into the objective function.

$$E(\phi) = \sum_{n=1}^{N} [y_n - \phi(x_n,1)]^T [y_n - \phi(x_n,1)] \qquad (2)$$

where $\phi(x,t)$ is a displacement field which is related to the velocity field by $\phi(x,1) = \int_0^1 v(\phi(x,t),t)dt$.

In [3], each source landmark traces a trajectory to approach the target landmark. The deformation energy is defined in terms of the smoothness of the flow field and the closeness of the landmark velocity to the time derivative of quasi-landmarks placed on the path connecting the two sets of landmarks

$$\begin{aligned} E(v,q) &= \int_0^1 \int_\Omega \parallel Lv(x,t) \parallel^2 dtdx \\ &+ \sum_{i=1}^{N} \int_0^1 \parallel \frac{dq_i(t)}{dt} - v(q_i(t),t) \parallel^2 dt \end{aligned} \quad (3)$$

where $q_i(t), \forall i \in \{1,\ldots,N\}$ are a set of quasi-landmarks lying on the path between the source and target landmarks. In [3], the authors construct a diffeomorphism between the two given sets of landmarks or labeled point-sets. They also formulated the cost function as a geodesic distance on the landmark manifold.

In [5], for the first time to our knowledge, the diffeomorphic matching approach was combined with a clustering objective to solve the problem of estimating a diffeomorphism between unlabeled point-sets. The objective function used combined a clustering cost, the deformation energy of geodesic spline and a coupling term between the two sets of cluster centers. In sharp contrast, the present work combines the landmark matching objective function (1) with a clustering objective to get a much simpler formulation than the one used in [5].

2 Joint clustering and diffeomorphism estimation

We have two sets of feature points $X = \{x_i | i = 1,2,...,N_1\}$ and $Y = \{y_j | j = 1,2,...,N_2\}$ in a 2D region Ω. The two point sets are to be clustered using a clustering objective. There are N clusters $\{r_k | k = 1,2,...,N\}$ for X and N clusters $\{s_k | k = 1,2,...,N\}$ for Y respectively. The cluster center indices are deemed to be in correspondence but the actual location of the clusters in each point-set is yet to be determined. A diffeomorphism from Ω to Ω maps the set of cluster centers $\{r_k | k = 1,2,...,N\}$ to $\{s_k | k = 1,2,...,N\}$ and we want this diffeomorphism to minimize a deformation energy. The overall objective function is

$$\begin{aligned} & E(M^x, M^y, r, s, v, \phi) \\ = & \sum_{i=1}^{N_1}\sum_{k=1}^{N} M_{ik}^x ||x_i - r_k||^2 + \sum_{j=1}^{N_2}\sum_{k=1}^{N} M_{jk}^y ||y_j - s_k||^2 \\ & + \sum_{k=1}^{N} ||s_k - \phi(r_k,1)||^2 + \lambda \int_0^1 \int_\Omega ||Lv(x,t)||^2 dxdt. \end{aligned} \quad (4)$$

In the above objective function, the cluster membership matrices satisfy $M_{ik}^x \in \{0,1\}, \forall ik$ $M_{jk}^y \in \{0,1\}, \forall jk$ and $\sum_{k=1}^N M_{ik}^x = 1, \sum_{k=1}^N M_{jk}^y = 1$. The matrix entry M_{ik}^x is the membership of data point x_i in cluster k whose center is at location r_k. The matrix entry M_{jk}^y is the membership of data point y_j in cluster k whose center is at position s_k. Again we discuss this as a pure clustering problem instead of in the context of mixture models. Consequently we have no information-theoretic restriction on the number of cluster centers N.

The diffeomorphic deformation energy in Ω is induced by the landmark displacements from r to s, where $x \in \Omega$ and $\phi(x,t)$ is a one parameter diffeomorphism: $\Omega \to \Omega$. Since the original point-sets differ in point count and are unlabeled, we cannot immediately use the diffeomorphism objective functions as in [2] or [3] respectively. Instead, the two point-sets are clustered and the landmark diffeomorphism objective is used between two sets of cluster centers r and s whose indices are always in correspondence. The diffeomorphism $\phi(x,t)$ is generated by the flow $v(x,t)$. $\phi(x,t)$ and $v(x,t)$ together satisfy the transport equation $\frac{\partial \phi(x,t)}{\partial t} = v(\phi(x,t),t)$ and the initial condition $\forall x$, $\phi(x,0) = x$ holds. This is in the inexact matching form and the displacement term $\sum_{k=1}^N ||s_k - \phi(r_k,1)||^2$ plays an important role here as the bridge between the two systems. This is also the reason why we prefer the deformation energy in this form because the coupling of the two sets of clusters appear naturally through the inexact matching term and we don't have to introduce external coupling terms as in [5]. Another advantage of this approach is that in this dynamic system described by the diffeomorphic group $\phi(x,t)$, the landmarks trace a trajectory exactly on the flow lines dictated by the field $v(x,t)$. Also, the feedback coupling is no longer needed as in the previous approach because with this deformation energy described above, if $\phi(x,t)$ is the minimizer of this energy, then $\phi^{-1}(x,t)$ is the backward mapping which also minimizes the same energy. This is formally described in the following theorem which is only valid for exact matching

Let $x_k \in \Omega$ and $y_k \in \Omega$, $k = 1,...,N$ and $\phi(x,t)$ is a one parameter diffeomorphism: $\Omega \to \Omega$ and it is generated by the flow $v(x,t)$. $\phi(x,t)$ and $v(x,t)$ together satisfy the transport equation $\frac{\partial \phi(x,t)}{\partial t} = v(\phi(x,t),t)$ and the initial condition $\forall x$, $\phi(x,0) = x$ holds.

Theorem. If $\phi(x_k, 1) = y_k$ and $\phi(x,t)$ and $v(x,t)$ minimize the energy $E = \int_0^1 \int_\Omega ||Lv(x,t)||^2 dxdt$, then the inverse mapping maps the landmarks backward $\phi^{-1}(y_k, 1) = x_k$ and $\phi^{-1}(x,t)$ and $-v(x,-t)$ also minimize the energy E.

Proof. First, from the known property of the diffeomorphism group of such a dynamical system, $\phi(x, t_1 + t_2) = \phi(\phi(x,t_1), t_2)$, it is easy to show that $\phi^{-1}(x,t) = \phi(x,-t)$. This is because $\phi(.,-t) \circ \phi(.,t)(x) = \phi(.,t) \circ \phi(.,-t)(x) = \phi(\phi(x,t),-t) = \phi(x, t + (-t)) = \phi(x,0) = x$. And $\phi(x,-t)$and $-v(x,-t)$ also satisfy the transport equation $\frac{\partial \phi(x,-t)}{\partial t} = -v(\phi(x,-t),-t)$. Suppose $\phi(x,t)$ and $v(x,t)$ minimize the energy $E = \int_0^1 \int_\Omega ||Lv(x,t)||^2 dxdt$, but $\phi^{-1}(x,t) = \phi(x,-t)$ and $-v(x,-t)$ *do not* minimize the energy $E = \int_0^1 \int_\Omega ||Lv(x,t)||^2 dxdt$. Let the minimizer be $\psi(x,t)$ and $u(x,t)$ such that $\forall k$, $\psi(y_k) = x_k$ and $\int_0^1 \int_\Omega ||Lu(x,t)||^2 dxdt < \int_0^1 \int_\Omega ||Lv(x,t)||^2 dxdt$. Then, we can construct $\psi^{-1}(x,t) = \psi(x,-t)$ such that $\psi^{-1}(x,t)$ and $-u(x,-t)$ satisfy the transport equation and $\psi^{-1}(x_k, 1) = y_k$. However $\int_0^1 \int_\Omega ||Lu(x,t)||^2 dxdt < \int_0^1 \int_\Omega ||Lv(x,t)||^2 dxdt$ contradicts the assumption that $v(x,t)$ is the minimizer of the energy E.

3 Algorithm summary

The joint clustering and diffeomorphism estimation algorithm has two components: i) diffeomorphism estimation and ii) clustering. For the diffeomorphism estimation, we expand the flow field in term of the kernel K of the L operator

$$v(x,t) = \sum_{k=1}^{N} \alpha_k(t) K(x, \phi_k(t)) \tag{5}$$

where $\phi_k(t)$ is notational shorthand for $\phi(r_k, t)$ and we also take into consideration the affine part of the mapping when we use thin-plate spline kernel with matrix entry $K_{ij} = r_{ij}^2 \log r_{ij}$ and $r_{ij} = \| x_i - x_j \|$. After discretizing in time t, the objective in (4) is expressed as

$$E = \sum_{i=1}^{N_1} \sum_{k=1}^{N} M_{ik}^x ||x_i - r_k||^2 + \sum_{j=1}^{N_2} \sum_{k=1}^{N} M_{jk}^y ||y_j - s_k||^2 \tag{6}$$

$$+ \sum_{k=1}^{N} \| s_k - r_k - \sum_{l=1}^{N} \sum_{t=0}^{S} [P(t) d_l(t) + \alpha_l(t) K(\phi_k(t), \phi_l(t))] \|^2$$

$$+ \lambda \sum_{k=1}^{N} \sum_{l=1}^{N} \sum_{t=0}^{S} < \alpha_k(t), \alpha_l(t) > K(\phi_k(t), \phi_l(t))$$

where

$$P(t) = \begin{pmatrix} 1 & \phi_1^1(t) & \phi_1^2(t) \\ . & . & . \\ . & . & . \\ . & . & . \\ 1 & \phi_N^1(t) & \phi_N^2(t) \end{pmatrix} \tag{7}$$

and d is the affine parameter matrix. After we perform a QR decomposition on P,

$$P(t) = (Q_1(t) : Q_2(t)) \begin{pmatrix} R(t) \\ 0 \end{pmatrix}. \tag{8}$$

We iteratively solve for $\alpha_k(t)$ and $\phi_k(t)$ using an alternating algorithm. When $\phi_k(t)$ is held fixed, we use the following approximation to solve for $\alpha_k(t)$. The solutions are

$$d(t) = R^{-1}(t) \left[Q_1(t) \phi(t+1) - Q_1(t) K(\phi(t)) Q_2(t) \gamma(t) \right] \tag{9}$$

$$\alpha(t) = Q_2(t) \gamma(t) \tag{10}$$

where $K(\phi(t))$ denotes the thin-plate spline kernel *matrix* evaluated at $\phi(t) \stackrel{\text{def}}{=} \{\phi(r_k, t) | k = 1, \ldots, N\}$ and

$$\gamma(t) = (Q_2^T(t) K(\phi(t)) Q_2(t) + \lambda)^{-1} Q_2^T(t) \phi(t+1). \tag{11}$$

When $\alpha_k(t)$ is held fixed, we use gradient descent to solve for $\phi_k(t)$:

$$\frac{\partial E}{\partial \phi_k(t)} = 2 \sum_{l=1}^{N} < \alpha_k(t), \alpha_l(t) - 2 W_l > \bigtriangledown_1 K(\phi_k(t), \phi_l(t)) \tag{12}$$

where $W_l = s_l - r_l - \sum_{m=1}^{N} \int_0^1 \alpha_m(t) K(\phi_m(t), \phi_l(t)) dt$.

The clustering of the two point-sets is handled by a deterministic annealing EM algorithm which iteratively estimates the cluster memberships M^x and M^y and the cluster centers r and s. The update of the memberships is the very standard E-step of the EM algorithm [6] and is performed as shown below.

$$M_{ik}^x = \frac{\exp(-\beta \|x_i - r_k\|^2)}{\sum_{l=1}^{C} \exp(-\beta \|x_i - r_l\|^2)}, \forall ik \tag{13}$$

$$M_{jk}^y = \frac{\exp(-\beta \|y_j - s_k\|^2)}{\sum_{l=1}^{C} \exp(-\beta \|y_j - s_l\|^2)}, \forall jk \tag{14}$$

where $\beta = \frac{1}{T}$ is the inverse temperature. The cluster center update is the M-step of the EM algorithm. This step is not the typical M-step. We use a closed-form solution for the cluster centers which is an approximation. From the

clustering standpoint, we assume that the change in the diffeomorphism at each iteration is *sufficiently small so that it can be neglected*. After making this approximation, we get

$$r_k = \frac{\sum_{i=1}^{N_1} M_{ik}^x x_k + s_k - \sum_{l=1}^{N} \int_0^1 \alpha_l(t) K(\phi_l(t), \phi_k(t)) dt}{1 + N_1}, \tag{15}$$

$$s_k = \frac{\sum_{j=1}^{N_2} M_{jk}^y y_j + \phi(r_k, 1)}{1 + N_2}, \forall k. \tag{16}$$

In the clustering and diffeomorphic estimations, we let λ vary proportionally with the temperature. This controls the rigidity of the mapping, starting from an almost rigid mapping while we obtain good correspondence and gradually softens so that good clustering is achieved. In this way both clustering and diffeomorphism are obtained simultaneously at convergence.

The overall algorithm is described below.

- **Initialization:** Initial temperature $T = 0.5(\max_i \|x_i - x_c\|^2 + \max_j \|y_j - y_c\|^2)$ where x_c and y_c are the centroids of X and Y respectively.
- **Begin A:** While $T > T_{\text{final}}$
  - **Step 1**: Clustering
  - Update memberships according to (13), (14).
  - Update cluster centers according to (15), (16).
  - **Step 2:** Diffeomorphism
  - Update (ϕ, v) by minimizing

$$\begin{aligned} E_{\text{diff}}(\phi, v) &= \sum_{k=1}^{C} ||s_k - \phi(r_k, 1)||^2 \\ &+ \lambda T \int_0^1 \int_\Omega ||Lv(x,t)||^2 dx dt \end{aligned}$$

  according to (9)(10) and (12).
- **Step 3:** Annealing. $T \leftarrow \gamma T$ where $\gamma < 1$.
- **End**

4 Results

We applied the algorithm to nine sets of 2D corpus callosum slices. The feature points were extracted with the help of a neuroanatomical expert. Figure 1 shows the nine corpus callosum 2D images, labeled CC1 through CC9. In our experiments, we first did the simultaneous clustering and matching with the corpus callosum point sets CC5 and CC9. The clustering of the two point sets is shown in Figure 2. There are 68 cluster centers. The circles represent the centers and the dots are the data points. The two cluster centers induce the diffeomorphic mapping of the 2D space. The warping of the 2D grid under this diffeomorphism is shown in Figure 3. Using this diffeomorphism, we calculated the after-image of original data points and compared them with the target data points. Due to the large number of cluster centers, the cluster centers nearly coincide with the original data points and the warping of the original data points is not shown in the figure. The correspondences (at the cluster level) are shown in Figure 4. The algorithm allows us to simultaneously obtain the diffeomorphism and the correspondence.

Using our formulation, we are able to calculate the geodesic distances between the two sets of cluster centers. This is done on the shape manifold. Each point q on the shape manifold M represents a set of N cluster centers $x_1, x_2, \ldots, x_N \in \mathbb{R}^2$ and has a coordinate $q = (x_1^1, x_1^2, x_2^1, x_2^2, ..., x_N^1, x_N^2)$ where $x_i = (x_i^1, x_i^2)$, $i = 1, 2, ..., N$. Let $q(t)$ be the geodesic path connecting two points q_1 and q_2 on the manifold. Using the norm defined for the tangent vector in [3], the geodesic distance between q_1 and q_2 is

$$D_{\text{geodesic}}(q_1, q_2) = \int_0^1 \sqrt{\lambda \dot{q}^T Q_2 (Q_2^T K(q) Q_2 + \lambda)^{-1} Q_2^T \dot{q}} dt. \tag{17}$$

where $K(q)$ is the kernel of the L operator evaluated at $q(t)$ and as mentioned previously, the thin-plate spline kernel is used. Q_2 comes from the QR decomposition of P

$$P = (Q_1 : Q_2) \begin{pmatrix} R \\ 0 \end{pmatrix} \tag{18}$$

and

$$P = \begin{pmatrix} 1 & x_1^1 & x_1^2 \\ . & . & . \\ . & . & . \\ . & . & . \\ 1 & x_N^1 & x_N^2 \end{pmatrix} \tag{19}$$

We also experimented with different number of cluster centers. Table 1 shows a modified Hausdorff distance as first introduced [7] between the image set of points CC5 after diffeomorphism and the target set of points CC9 when the number of clusters vary. The reason for using the modified Hausdorff distance instead of the Hausdorff distance is that the latter is too sensitive to outliers. The definition of the modified Hausdorff distance is

$$H_{\text{mod}}(A, B) = \max(h_{\text{mod}}(A, B), h_{\text{mod}}(B, A)), \tag{20}$$

where A and B are finite point sets and

$$h_{\text{mod}}(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} \| a - b \| \tag{21}$$

is the average of the minimum distances instead of the maximum of the minimum distances. It is easy to see that when the number of clusters increases, the matching improves as the modified Hausdorff distance decreases.

In the third column in Table 2, we list the geodesic distances between the two sets of cluster centers after pairwise warping and clustering all the pair of corpus callosum point sets. Using the cluster centers as landmarks, a diffeomorphic mapping of the space is induced. With this induced diffeomorphism, we mapped the original data sets and compared the image of the original point set under the diffeomorphism and the target point set using the modified Hausdorff distance. The modified Hausdorff distances between the pairs are listed in the fourth column in Table 2. Finally, from the original nine corpus callosum point sets, we warped the first eight point sets onto the ninth set and Figure 5 displays the overlay of all point sets after diffeomorphic warping.

5 Discussion

We have designed an objective function and an algorithm to simultaneously find the best clustering of two point-sets and a mapping with the least deformation of space. We require the space deformation to be a diffeomorphic mapping because it is smooth and homeomorphic at the same time. The essence of this requirement is that in a homeomorphic mapping, neighboring points are mapped to neighboring points and the same is true for the inverse mapping. A diffeomorphic mapping can preserve the features of shapes. The diffeomorphism parameterization allows us to recover large deformations while simultaneously achieving good correspondence. After diffeomorphism estimation, the shape distance, defined as the geodesic distance on the shape manifold is computed. Since the point-sets have different cardinalities and since the shape distance is only defined w.r.t. the cluster centers, we also computed a modified Hausdorff distance between one original point-set and the after-image of a second point-set. We conclude that when the number of cluster centers increases, the modified Hausdorff distance decreases. This means that the original data points are better represented by the cluster centers. Since this only holds in less noisy situations, further work is needed to fix the number of cluster centers and the regularization parameter when operating with noisy point-sets.

References

[1] C. Small, *The statistical theory of shape*. New York, NY: Springer, 1996.

[2] S. Joshi and M. Miller, "Landmark matching via large deformation diffeomorphisms," *IEEE Trans. Image Processing*, vol. 9, pp. 1357–1370, 2000.

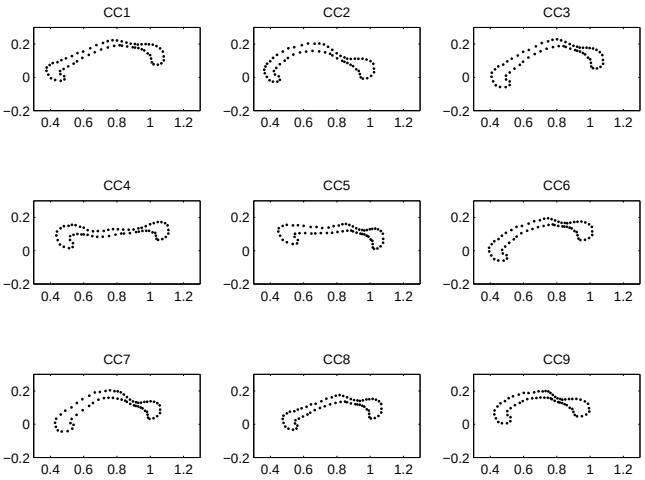

Figure 1: Point sets of nine corpus callosum images.

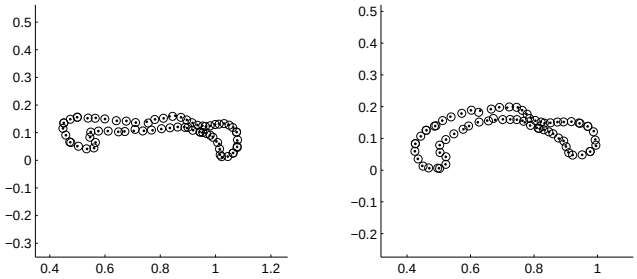

Figure 2: Clustering of the two point sets.

[3] V. Camion and L. Younes, "Geodesic interpolating splines," in *Energy Minimization Methods for Computer Vision and Pattern Recognition*, pp. 513–527, New York: Springer, 2001.

[4] G. Wahba, *Spline models for observational data*. Philadelphia, PA: SIAM, 1990.

[5]

[6]

[7] M. P. Dubuisson and A. K. Jain, "A modifi ed hausdorff distance for object matching," *ICPR94*, pp. A:566–568, 1994.

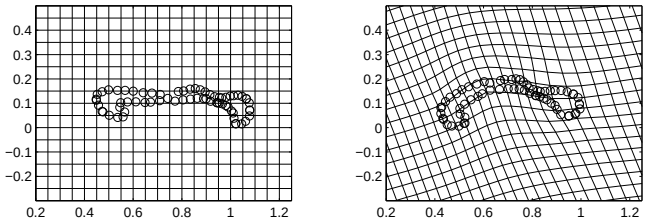

Figure 3: Diffeomorphic mapping of the space.

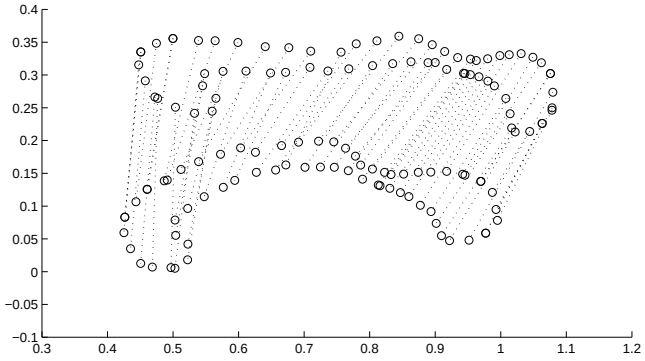

Figure 4: Matching between the two point sets.

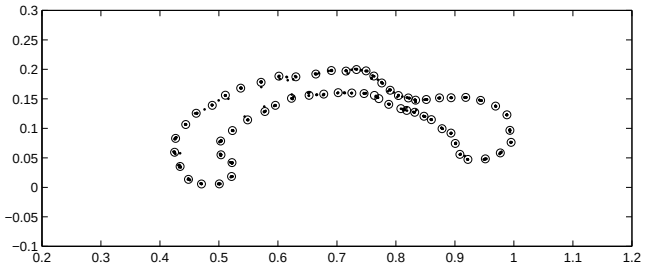

Figure 5: Overlay of the after-images of eight point sets with the ninth set.

Number of Clusters	Modified Hausdorff Distance
10	0.0082
20	0.0082
30	0.0057
40	0.0050
50	0.0043
60	0.0035
68	0.0027

Table 1: Modified Hausdorff distance of the matching point sets.

From	To	Geodesic distance	M.H. distance
CC1	CC2	0.0264	0.0055
CC1	CC3	0.0132	0.0014
CC1	CC4	0.0289	0.0048
CC1	CC5	0.0269	0.0056
CC1	CC6	0.0250	0.0097
CC1	CC7	0.0323	0.0054
CC1	CC8	0.0256	0.0043
CC1	CC9	0.0241	0.0041
CC2	CC3	0.0277	0.0059
CC2	CC4	0.0342	0.0063
CC2	CC5	0.0308	0.0057
CC2	CC6	0.0211	0.0100
CC2	CC7	0.0215	0.0040
CC2	CC8	0.0271	0.0044
CC2	CC9	0.0258	0.0093
CC3	CC4	0.0443	0.0059
CC3	CC5	0.0294	0.0047
CC3	CC6	0.0181	0.0032
CC3	CC7	0.0256	0.0060
CC3	CC8	0.0153	0.0018
CC3	CC9	0.0305	0.0044
CC4	CC5	0.0231	0.0046
CC4	CC6	0.0304	0.0056
CC4	CC7	0.0324	0.0056
CC4	CC8	0.0311	0.0054
CC4	CC9	0.0434	0.0090
CC5	CC6	0.0266	0.0056
CC5	CC7	0.0325	0.0053
CC5	CC8	0.0225	0.0037
CC5	CC9	0.0305	0.0069
CC6	CC7	0.0244	0.0050
CC6	CC8	0.0186	0.0026
CC6	CC9	0.0274	0.0056
CC7	CC8	0.0212	0.0044
CC7	CC9	0.0241	0.0102
CC8	CC9	0.0196	0.0050

Table 2: Geodesic distances between two sets of cluster centers and modified Hausdorff distances of matching points.