A Hierarchical Framework for Spectral Correspondence

Marco Carcassoni and Edwin R. Hancock

Department of Computer Science,
University of York,York YO1 5DD, UK.
{marco,erh}@cs.york.ac.uk

Abstract. The modal correspondence method of Shapiro and Brady aims to match point-sets by comparing the eigenvectors of a pairwise point proximity matrix. Although elegant by means of its matrix representation, the method is notoriously susceptible to differences in the relational structure of the point-sets under consideration. In this paper we demonstrate how the method can be rendered robust to structural differences by adopting a hierarchical approach. We place the modal matching problem in a probabilistic setting in which the correspondences between pairwise clusters can be used to constrain the individual point correspondences. To meet this goal we commence by describing an iterative method which can be applied to the point proximity matrix to identify the locations of pairwise modal clusters. Once we have assigned points to clusters, we compute within-cluster and between-cluster proximity matrices. The modal co-efficients for these two sets of proximity matrices are used to compute cluster correspondence and cluster-conditional point correspondence probabilities. A sensitivity study on synthetic point-sets reveals that the method is considerably more robust than the conventional method to clutter or point-set contamination.

1 Introduction

Eigendecomposition, or modal analysis, has proved to be an alluring yet elusive method for correspondence matching. Stated simply, the aim is to find the pattern of correspondence matches between two sets of objects using the eigenvectors of an adjacency matrix or an attribute proximity matrix. The problem has much in common with spectral graph theory [1] and has been extensively studied for both the abstract problem of graph-matching [17,16], and for point pattern matching [14,12,11]. In the case of graph-matching the adjacency matrix represents either the weighted or unweighted edges of the relational structure under study. For point pattern matching, the proximity matrix represents the pairwise distance relationships. The method may be implemented in a number of ways. The simplest of these is to minimize the distance between the modal coefficients. A more sophisticated approach is to use a factorization method such as singular value decomposition to find the permutation matrix which minimizes the differences between the adjacency structures. Unfortunately, the method invariable fails when the sets of objects being matched are not of the same size due

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2350, pp. 266–281, 2002.

to structural differences. The reason for this is that the pattern of eigenvectors is unstable when structural differences are present.

There are several concrete examples in the pattern analysis literature. Turning our attention to graph-matching, Umeyama has an eigendecomposition method that recovers the permutation matrix that maximizes the correlation or overlap of the adjacency matrices for graphs of the same size [17]. This method uses a factorization method to find the permutation matrix that brings the two graphs into correspondence. Horaud and Sossa [5] have adopted a purely structural approach to the recognition of line-drawings. Their representation is based on the immanental polynomials for the Laplacian matrix of the line-connectivity graph. By comparing the coefficients of the polynomials, they are able to index into a large data-base of line-drawings. Of more direct relevance to this paper is the literature on point-pattern matching. Borrowing ideas from structural chemistry, Scott and Longuet-Higgins were among the first to use eigendecomposition methods for point correspondence analysis [12]. They showed how to recover correspondences via singular value decomposition on the point association matrix between different images. However, the method has a number of well documented problems relating to the small range of scale and angle differences for which it is effective. In an attempt to overcome these problems, Shapiro and Brady [14] have developed a method in which point sets are matched by comparing the eigenvectors of the point proximity matrix. Here the proximity matrix is constructed by computing the Gaussian weighted distance between points. Matching between different point-sets is effected by comparing the pattern of eigenvectors, or modal co-efficients, in different images. The method extends the range of angle and scale differences over which reliable correspondences can be recovered. However, the method fails for point sets of different sizes. In a recent paper [2] we have revisited the method of Shapiro and Brady. Our aim was to use the correspondence information delivered by the method to develop an EM algorithm for point-set alignment. For structurally intact point-sets subject to positional jitter, we showed that the performance of the Shapiro and Brady method could be improved using ideas from robust statistics to compute the proximity matrix and to compare the modal co-efficients. To overcome the difficulties encountered with point-sets of different size, an explicit alignment process was required.

The aim in this paper is to return to the Shapiro and Brady [14] method and to focus on how the method can be rendered robust to structural differences in the point-sets. We adopt a hierarchical approach. The method is based on the observation that the modes of the proximity matrix can be viewed as pairwise clusters. Moreover, the modal co-efficients represent the affinity of the raw points to the clusters. This idea has been exploited by several authors to develop powerful image segmentation [15] and grouping methods [13,8,6]. Sengupta and Boyer [13] have used property matrix spectra to characterise line-patterns. Various attribute representations are suggested and compared. Shokoufandeh, Dickinson and Siddiqi [16] have shown how graphs can be encoded using local topological spectra for shape recognition from large data-bases. Sarkar and Soundararajan [10] have shown how graph-spectral methods can be combined with cellular automata to learn grouping structure. Finally, a number of authors

have used spectral methods to perform pairwise clustering on image data. Shi and Malik [15] use the second eigenvalue to segment images by performing an eigen-decomposition on a matrix of pairwise attribute differences using the iterative normalised cut method. Sarkar and Boyer [9] and Freeman and Perona [8] have developed matrix factorisation methods for line-grouping. In a recent paper, Weiss [18] has compared a number of matrix factorisation methods for matching and segmentation, and has commented on the way in which they compute the proximity matrix. His conclusion was that performance could be significantly improved if the matrix is correctly normalised. Inoue and Urahama [6] have shown how the sequential extraction of eigen-modes can be used to cluster pairwise pixel data as an alternative to computationally expensive methods, such as the mean-field annealing idea of Buhmann and Hoffman [4]. Rather than explicitly grouping the points prior to matching, here we aim to characterise the potential groupings in an implicit or probabilistic way and to exploit their arrangement to provide constraints on the pattern of correspondences.

Our approach is as follows. Each mode of the point-proximity matrix is taken to represent a potential grouping or cluster of points. For each group, we can compute a cluster centre point-position. While the pattern of modal co-efficients of the proximity matrix may be disturbed by structural differences in the point-sets, the centre-points of the groups or clusters may be more stable. Hence, we can use the cluster-centre proximity matrix to improve the correspondence process. Here we use an evidence combining method which is posed in a hierarchical framework. We compute the probability that pairs of points are in correspondence by developing a mixture model over the set of possible correspondences between the most significant groupings of points. In this way the cluster-centre correspondences, weight the point-correspondence probabilities. We discuss various alternative ways in which the correspondence process may be modelled using the modal co-efficients of the point and cluster centre proximity matrices. We compare these alternatives with both the Shapiro and Brady method and our previously reported method.

2 Point Correspondence

The modal approach to point correspondence introduced by Shapiro and Brady [14] commences by enumerating a point proximity matrix. This is a continuous or weighted counterpart of the graph adjacency matrix. Rather than setting the elements to unity or zero depending on whether or not there is a connecting edge between a pair of nodes, the elements of the proximity matrix are weights that reflect the strength of a pairwise adjacency relation. The weights of the proximity matrix are computed by taking a Gaussian function of the interpoint distances, Once the proximity matrix is to hand, then correspondences are located by computing its eigenvectors. The eigenvectors of the proximity matrix become the columns of a transformation matrix which operates on the original point identities. The rows of the transformation matrix represent the components of the original points in the directions of the eigenvectors. We can locate point

correspondences by searching for rows of the transformation matrix which have maximal similarity.

Unfortunately there are two drawbacks with this modal method of correspondence. Firstly, there is no clear reason to use Gaussian weighting in favour of possible alternatives. Gaussian weighting may not be the most suitable choice to control the effects of pattern distortion due to point movement under measurement error or deformation under affine or perspective geometry. Secondly, the method proves fragile to structural differences introduced by the addition of clutter or point drop-out. In a recent paper we have addressed the first of these problems by using robust error kernels to compute the proximity matrix [2]. Here we focus on the second problem, and develop a hierarchical method matching point-sets.

In this section we review the existing work on the modal matching of point-sets, before detailing an improved method aimed at overcoming the problem of different point set size.

2.1 Shapiro and Brady

We are interested in finding the correspondences between two point-sets, a model point-set $\mathbf{z}$ and a data point-set $\mathbf{w}$. Each point in the image data set is represented by an position vector co-ordinates $\underline{\mathrm{w}}_i = (x_i, y_i)^T$ where i is the point index. In the interests of brevity we will denote the entire set of image points by $\mathbf{w} = \{\underline{\mathrm{w}}_i, \forall i \in \mathcal{D}\}$ where $\mathcal{D}$ is the point set. The corresponding fiducial points constituting the model are similarly represented by $\mathbf{z} = \{\underline{\mathrm{z}}_j, \forall j \in \mathcal{M}\}$ where $\mathcal{M}$ denotes the index-set for the model feature-points $\underline{\mathrm{z}}_j$.

The role of the weighting function used to compute the elements of the proximity matrix is to model the probability of adjacency relations between points. The standard way to represent the adjacency relations between points is to use the Gaussian proximity matrix. If i and i' are two data points, then the corresponding element of the proximity matrix is given by

$$H_D(i, i') = \exp\left[-\frac{1}{2s^2}||\underline{\mathrm{w}}_i - \underline{\mathrm{w}}_{i'}||^2\right] \tag{1}$$

The modal structure of the two point-sets is found by solving the eigenvalue equation $det[H - \lambda I] = 0$ together with the associated eigenvector equation $H\phi_l = \lambda_l\phi_l$, where λ_l is the l^{th} eigenvalue of the matrix H and ϕ_l is the corresponding eigenvector. The vectors are ordered according to the magnitude of the associated eigenvalues. The ordered column-vectors are used to construct a modal matrix $\Phi = (\phi_1|\phi_2|\phi_3|.....)$. The column index of this matrix refers to the magnitude order of the eigenvalues while the row-index is the index of the original point-set. This modal decomposition is repeated for both the data and transformed model point-sets to give a data-point modal matrix $\Phi_D = (\phi_1^D|\phi_2^D|\phi_3^D|...|\phi_{|\mathcal{D}|}^D)$ and a model-point modal matrix $\Phi_M = (\phi_1^M|\phi_2^M|\phi_3^M|...|\phi_{|\mathcal{M}|}^M)$. Since the two point-sets are potentially of different size, the modes are truncated of the larger point-set. This corresponds to removing the last $||\mathcal{D}| - |\mathcal{M}||$ rows and columns of the larger matrix. The resulting matrix has $o = \min[\mathcal{D}, \mathcal{M}]$ rows and columns.

The modal matrices can be viewed as inducing a linear transformation on the original identities of the point-sets. Each row of the modal matrix represents one of the original points. The column entries in each row measure how the original point identities are distributed among the different eigen-modes.

Based on this eigendecomposition Shapiro and Brady [14] find correspondences by comparing the rows of the model matrices Φ_M and Φ_D. The decision concerning the correspondences is made on the basis of the similarity of different rows in the modal matrices for the data and the model. The measure of similarity is the Euclidean distance between the elements in the corresponding rows. According to Shapiro and Brady the correspondence probabilities are assigned according to the following binary decision

$$\zeta_{i,j}^{SB} = \begin{cases} 1 & \text{if } j = \arg\min_{j'} \sum_{l=1}^{o} ||\Phi_D(i,l) - \Phi_M(j',l)||^2 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The decision regarding the most likely correspondence can then be made on the basis of the maximum value of the probability.

2.2 Prior Work

In this section we briefly review our previous work aimed at improving the modal matching method. It must be stressed that the aim of this work was to compute correspondence probabilities for the purposes of point-set alignment using a variant of the EM algorithm. For point-sets of the same size which were not subject to contamination by clutter or dropout, we showed that the accuracy of correspondence matching could be improved by a) using a weighting function suggested by robust statistics to compute the point proximity matrix and b) comparing the modal co-efficients using a robust statistical procedure.

In Shapiro and Brady's original work the weighting function was the Gaussian [14]. Our first contribution has been to show that alternative weighting functions suggested by the robust statistics literature offer significant improvements [2].

According to robust statistics, there are some choices of possible weighting functions. In our previous work [2] we showed that the sigmoidal weighting function, generated by the hyperbolic tangent function

$$H_D(i,i') = \frac{2}{\pi||\underline{w}_i - \underline{w}_{i'}||} \log\cosh\left[\frac{\pi}{s}||\underline{w}_i - w_{i'}||\right] \tag{3}$$

gives improved performance under positional jitter.

The second contribution was to show that the method of assigning correspondences could be significantly improved if the elements of the modal matrix were compared using a robust statistical procedure. When there is a significant difference between one or more of the components of the eigenvectors, then these errors dominate the Euclidean distance measure used by Shapiro and Brady. One way to make the computation of correspondences robust to outlier measurement error is to accumulate probability on a component by component basis over the eigenvectors. To do this assume that the individual elements of the modal matrix

are subject to Gaussian measurement errors and compute the correspondence probability using the formula

$$\zeta_{i,j}^{CH} = \frac{\sum_{l=1}^{o} \exp\left[-k||\Phi_D(i,l) - \Phi_M(j,l)||^2\right]}{\sum_{j' \in \mathcal{M}} \sum_{l=1}^{o} \exp\left[-k||\Phi_D(i,l) - \Phi_M(j',l)||^2\right]} \tag{4}$$

where k is a constant. In this way large measurement errors contribute insignificantly through the individual exponentials appearing under the summation over the components of the eigenvectors.

These two refinements of Shapiro and Brady's method offer some improvements in terms of robustness to positional jitter and affine skew. However, when the point-sets under study are of different size, i.e. they are subject to structural corruption, then both methods fail. The reason for this is that the co-efficients of the modal matrices become unstable and can not be used for correspondence matching. Our aim in to this paper is to suggest a way of overcoming this problem.

3 Modal Clusters

Our aim if to pose the modal correspondence of point-sets in a hierarchical framework. We commence by locating the modal clusters of the point-sets under study. This is an iterative process which alternates between two steps. The first step involves computing the mean position vector for each mode of the proximity matrix. The second step involves computing a revised proximity matrix from the mean modal position vectors. Once this iterative process has converged, we select the mean position vectors associated with the most significant modes of the proximity matrix. These position-vectors are used to compute a modal-cluster proximity matrix. By using constraints provided by the modal correspondences of the cluster-centres, we aim to improve the correspondence matching of the raw point-sets. In this section, we describe how to perform the iterative modal clustering and how to compute the modal proximity matrix.

The coefficients of the modal matrix Φ can be viewed as providing information concerning pairwise clusters of points. Each mode, i.e. each column of the modal matrix Φ_D, is represented by an orthogonal vector in a $|\mathcal{D}|$ dimensional space. The columns associated with the eigenvalues of largest magnitude represent the most significant arrangements of points, while those associated with the eigenvalues of smallest magnitude represent insignificant structure. For a given point i the different modal co-efficients $\Phi(i,l)$, $l = 1, ..., |\mathcal{D}|$ represent the affinity of the point to the different clusters. The larger the magnitude of the co-efficient, the greater the cluster affinity. In other words, the entries in the columns of the modal matrix represent the membership affinities for the different clusters. The row-entries, on the other hand represent the way in which the individual points are distributed among the different clusters. Here we aim to exploit this property of the modal matrix to develop a fast and robust matching method.

Our idea is based on the simple observation, that while the modal coefficients, i.e. the entries in the columns of the modal matrix, may not be stable under the addition of extra points, the physical centre of the associated cluster will be relatively robust to the addition of outliers.

3.1 Iterative Computation of the Modes

To locate the cluster-centres we adopt an iterative process. At each iteration, we use the modal co-efficients to compute a mean position vector for each eigen-mode. These modal-centres are then used to compute a revised proximity matrix. In their turn, the modal co-efficients for this revised proximity matrix are used to update the positions of the modal centres.

To this end we compute a mean position-vector for each eigen-mode. Let $\Phi_D^{(n)}$ be the modal matrix at iteration n. For the mode with eigenvalue λ_l, the position-vector for the cluster centre is

$$\underline{c}_l^{D\ (n)} = \frac{\sum_{i=1}^{|\mathcal{D}|} |\Phi_D^{(n)}(i,l)| \underline{w}_i}{\sum_{i=1}^{|\mathcal{D}|} \Phi_D^{(n)}(i,l)|} \tag{5}$$

Next, we compute the revised proximity matrix for the modal position vectors. The elements of the proximity matrix are again computed using the robust weighting kernel and are given by

$$H_D^{(n)}(l,l') = \frac{2}{\pi ||\underline{c}_l^{D\ (n)} - \underline{c}_{l'}^{D\ (n)}||} \log\cosh\left[\frac{\pi}{s}||\underline{c}_l^{D\ (n)} - \underline{c}_{l'}^{D\ (n)}||\right] \tag{6}$$

By solving the eigenvalue equation $det[H_D^{(n)} - \lambda^{(n)} I] = 0$ together with the associated eigenvector equation $H_D^{(n)}\phi_l^{(n)} = \lambda_l^{(n)}\phi_l^{(n)}$, we compute the updated a modal matrix $\Phi^{(n)} = (\phi_1^{(n)}|\phi_2^{(n)}|\phi_3^{(n)}|.....)$. This process is iterated until the modal position vectors stabilize. The final modal position vectors are noted by $\underline{c}_l^{D\ (\infty)}$ and the final modal co-efficient matrix by $\Phi_D^{(\infty)}$.

Once the pairwise clustering process has converged, then we can assign points to modal clusters. We represent the arrangement of points using both a set of within-cluster proximity matrices and a single between-cluster proximity matrixes. The modal structure of the between-cluster proximity matrix is used to compute the probabilities that individual cluster centres are in correspondence. The modal structure of the within-cluster proximity matrices are used to compute the probability that individual points within corresponding clusters match to one-another. Details of how the within-cluster and between-cluster modal structure is computed are outlined in the subsequent two subsections of the paper.

3.2 Within-Cluster Modal Matrices

When the iterative clustering process has converged, then the elements of the modal matrix can be used to assign points to clusters. We are interested in

using the modal co-efficients and the cluster centre locations to compute the probability $P(i \in \omega_d)$ that the node i belongs to the cluster associated with mode ω_d of the original point-set. We use the co-efficients of the first S columns of the modal matrix $\Phi_D^{(\infty)}$ to compute the cluster membership probability. Here we assume that cluster membership probability is proportional to the magnitude of the entry in the row indexed i and column indexed ω_d of the modal matrix $\Phi_D^{(\infty)}$ and write

$$P(i \in \omega_d) = \Phi_D^*(i, \omega_d) = \frac{|\Phi_D^{(\infty)}(i, \omega_d)|}{\sum_{l=1}^{S} |\Phi_D^{(\infty)}(i, \omega_d)|} \tag{7}$$

For the points belonging to each cluster, we construct a within-cluster proximity matrix. To do this we first identify the points which belong to each modal cluster. This is done of the basis of the cluster-membership probabilities $P(i \in \omega_d)$. The set of points assigned to the cluster ω_D is $\mathcal{C}_{\omega_d}^D = \{i | P(i \in \omega_d) > T_c\}$ where T_c is a membership probability threshold. To construct this matrix we will need to relabel the points using a cluster point index which runs from 1 to $|\mathcal{C}_{\omega_d}|$. Accordingly we let δ_{i,ω_d}^D denote the point-index assigned to the node i in the cluster ω_d. The proximity matrix for the points belonging to this cluster is denoted by F_{ω_D} and the corresponding modal matrix is $\Theta_{\omega_d}^D$. The modal matrix for the cluster indexed ω_m in the model point-set is denoted by $\Theta_{\omega_m}^M$.

3.3 Between Cluster Modal Matrix

We also construct a between-cluster modal matrix to summarize the global structure or arrangement of the original point-set $\mathbf{w}$. To do this we select the positions of the cluster-centres for the S largest eigenvalues, i.e. the first S columns of $\Phi_D^{(\infty)}$. There are a number of ways of choosing S. Here we set the value of S so that the co-efficients of the subsequent columns of $\Phi_D^{(\infty)}$ are insignificant. If T is a threshold, then the condition is that $|\Phi_D^{(\infty)}(i, l)| < T$ for $i = 1, ..., |\mathcal{D}|$ and $l > S$. Our idea is to use the modes of the $S \times S$ cluster-centre proximity matrix G_D for the purposes of matching. Accordingly, we solve the equation $det(G_D - \Lambda^D I) = 0$ to locate the eigenvalues of the modal cluster-centre proximity matrix. The eigenvectors ψ_L, $L = 1, .., S$ of the cluster-centre proximity matrix are found by solving the equation $G_D \psi_l^D = \Lambda_l^D \psi_l^D$ As before, these eigenvectors can be used to construct a modal-matrix for the cluster centre positions. The matrix has the eigenvectors of G as columns, i.e. $\Psi_D = \left(\psi_1^D | \psi_2^D | ..... \psi_S^D\right)$ This procedure is repeated to construct a second $S \times S$ cluster-centre modal matrix Ψ_M for the set of model points $\mathbf{z}$. Since the principal modal-clusters are selected on the magnitude-order of the associated eigenvalues, there is no need to re-order them.

4 Matching

The aim in this paper is to explore whether the additional information provided by the modal clusters can be used to improve the robustness of the matching

process to point addition and dropout. We would like to compute the probability $P(i \leftrightarrow j)$, that the data-point $i \in \mathcal{D}$ is in correspondence with the model data-point $j \in \mathcal{M}$. To do this we construct a mixture model over the set of possible correspondences between the set of S modal clusters extracted from the data point positions and the model point positions. Suppose that ω_d and ω_m respectively represent labels assigned to the modal clusters of the data and model point-sets. Applying the Bayes formula, we can write

$$P(i \leftrightarrow j) = \sum_{\omega_d=1}^{S} \sum_{\omega_m=1}^{S} P(i \leftrightarrow j|\omega_d \leftrightarrow \omega_m) P(\omega_d \leftrightarrow \omega_m) \tag{8}$$

where $P(i \leftrightarrow j|\omega_d \leftrightarrow \omega_m)$ represents the cluster-conditional probability that the node i belonging to the data-graph cluster ω_d is in correspondence with the node j that belongs to the model-graph cluster ω_m. The quantity $P(\omega_d \leftrightarrow \omega_m)$ denotes the probability that the data point-set cluster indexed ω_d is in correspondence with the model point-set cluster indexed ω_m.

4.1 Cluster Conditional Correspondence Probabilities

To compute the cluster-conditional point correspondence probabilities we use the modal structure of the within-cluster proximity matrices. These correspondence probabilities are computed using the method outlined in Equation (4). As a result, we write

$$\begin{aligned} &P(i \leftrightarrow j|\omega_d \leftrightarrow \omega_m) = \\ &= \frac{\sum_{l=1}^{O_{\omega_d,\omega_m}} \exp\left[-k_w||\Theta^D_{\omega_d}(\delta^D_{i,\omega_d}, l) - \Theta^M_{\omega_m}(\delta^D_{j\omega_m}, l)||^2\right]}{\sum_{j' \in \mathcal{M}} \sum_{l=1}^{O_{\omega_d,\omega_m}} \exp\left[-k_w||\Theta^D_{\omega_d}(\delta^D_{i,\omega_d}, l) - \Theta^M_{\omega_m}(\delta^M_{j',\omega_m}, l)||^2\right]} \end{aligned} \tag{9}$$

where $O_{\omega_d,\omega_m} = min[|\mathcal{C}_{\omega_m}|, |\mathcal{C}_{\omega_d}||$ is the size of the smaller cluster.

4.2 Cluster Correspondence Probabilities

We have investigated two methods for computing the cluster correspondence probabilities $P(\omega_d \leftrightarrow \omega_m)$:

– **Modal eigenvalues:** The first method used to compute the cluster-centre correspondence probabilities relies on the similarity of the normalized eigenvalues of the cluster-centre modal matrix. The probabilities are computed in the following manner

$$\begin{aligned} &P(\omega_d \leftrightarrow \omega_m) = \\ &\frac{\exp\left[-k_e\left\{\frac{|\Lambda^D_{\omega_d}|}{\sum_{\omega_d=1}^{S}|\Lambda^D_{\omega_d}|} - \frac{|\Lambda^M_{\omega_m}|}{\sum_{\omega_m=1}^{S}|\Lambda^M_{\omega_m}|}\right\}^2\right]}{\sum_{\omega_m=1}^{S}\exp\left[-k_e\left\{\frac{|\Lambda^D_{\omega_d}|}{\sum_{\omega_d=1}^{S}|\Lambda^D_{\omega_d}|} - \frac{|\Lambda^M_{\omega_m}|}{\sum_{\omega_m=1}^{S}|\Lambda^M_{\omega_m}|}\right\}^2\right]} \end{aligned} \tag{10}$$

– **Modal co-efficients:** The mode correspondence probabilities have also been computed by performing a robust comparison of the co-efficients of the modal matrices of the cluster-centre proximity matrix. This is simply an application of the method outlined in Equation (4) to the modal co-efficients of the between-cluster proximity matrix. We therefore set

$$P(\omega_d \leftrightarrow \omega_m) = \frac{\sum_{L=1}^{S} \exp\left[-k_b||\Psi_D(\omega_d, L)| - |\Psi_M(\omega_m, L)||^2\right]}{\sum_{\omega_m=1}^{S}\sum_{L=1}^{S} \exp\left[-k_b||\Psi_D(\omega_d, L)| - |\Psi_M(\omega_m, L)||^2\right]} \tag{11}$$

Note that we no-longer have to truncate the number of modes of the larger point-set since we have chosen only the S principal clusters from both the model and data.

4.3 Correspondence Probabilities

Using these models for the within and between-cluster modal co-efficients we develop two models for the correspondence probabilities appearing in Equation (10):

– Modal co-efficients

$$P(i \leftrightarrow j) = \sum_{\omega_d=1}^{S}\sum_{\omega_m=1}^{S} \frac{\sum_{l=1}^{O_{\omega_d,\omega_m}} \exp\left[-k_w||\Theta^D_{\omega_d}(\delta^D_{i,\omega_d}, l) - \Theta^M_{\omega_m}(\delta^D_{j\omega_m}, l)||^2\right]}{\sum_{j'\in\mathcal{M}}\sum_{l=1}^{O_{\omega_d,\omega_m}} \exp\left[-k_w||\Theta^D_{\omega_d}(\delta^D_{i,\omega_d}, l) - \Theta^M_{\omega_m}(\delta^M_{j',\omega_m}, l)||^2\right]} \frac{\sum_{L=1}^{S} \exp\left[-k_b||\Psi_D(\omega_d, L)| - |\Psi_M(\omega_m, L)||^2\right]}{\sum_{\omega_m=1}^{S}\sum_{L=1}^{S} \exp\left[-k_b||\Psi_D(\omega_d, L)| - |\Psi_M(\omega_m, L)||^2\right]} \tag{12}$$

According to this formula, the correspondence match between the points i and j receives support if they belong to clusters which have a high probability of modal correspondence.

– Using eigenvectors and eigenvalues

$$P(i \leftrightarrow j) = \sum_{\omega_d=1}^{S}\sum_{\omega_m=1}^{S} \frac{\sum_{l=1}^{O_{\omega_d,\omega_m}} \exp\left[-k_w||\Theta^D_{\omega_d}(\delta^D_{i,\omega_d}, l) - \Theta^M_{\omega_m}(\delta^D_{j\omega_m}, l)||^2\right]}{\sum_{j'\in\mathcal{M}}\sum_{l=1}^{O_{\omega_d,\omega_m}} \exp\left[-k_w||\Theta^D_{\omega_d}(\delta^D_{i,\omega_d}, l) - \Theta^M_{\omega_m}(\delta^M_{j',\omega_m}, l)||^2\right]} \tag{13}$$

$$\frac{\sum_{L=1}^{S}\exp\left[-k_b||\Psi_D(\omega_d,L)|-|\Psi_M(\omega_m,L)||^2\right]}{\sum_{\omega_m=1}^{S}\sum_{L=1}^{S}\exp\left[-k_b||\Psi_D(\omega_d,L)|-|\Psi_M(\omega_m,L)||^2\right]}$$

$$\frac{\exp\left[-k_e\left\{\frac{|\Lambda^D_{\omega_d}|}{\sum_{\omega_d=1}^{S}|\Lambda^D_{\omega_d}|}-\frac{|\Lambda^M_{\omega_m}|}{\sum_{\omega_m=1}^{S}|\Lambda^M_{\omega_m}|}\right\}^2\right]}{\sum_{\omega_m=1}^{S}\exp\left[-k_e\left\{\frac{|\Lambda^D_{\omega_d}|}{\sum_{\omega_d=1}^{S}|\Lambda^D_{\omega_d}|}-\frac{|\Lambda^M_{\omega_m}|}{\sum_{\omega_m=1}^{S}|\Lambda^M_{\omega_m}|}\right\}^2\right]}$$

4.4 Parameters

There are a number of parameters which need to be controlled in our correspondence matching method. The first of these are the widths s of the error kernels used to compute the proximity matrices for the original points, the individual clusters and the cluster-centres. In each case, we have found that the best method is to set s to be equal to the median interpoint distance. There are three exponential constants k_e, k_b and k_w which must be set for the correspondence probability computations. Here we set all three constants to 0.1.

5 Experiments

In this section we describe our experimental evaluation of the new modal correspondence method. This is divided into two parts. We commence with a sensitivity study on synthetic data. This is aimed at measuring the effectiveness of the method when the point sets under study are subject to clutter and positional jitter. The second part of the study focuses on real world data. Here we investigate the method when applied to finding point correspondences between curvature features in gesture sequences.

5.1 Sensitivity Study

In our sensitivity study, we have compared the new correspondence method with those of Shapiro and Brady [14] and our previous work [2]. The Shapiro and Brady method is based purely on modal correspondence analysis, while the our previous method uses modal correspondence probabilities to weight the estimation of affine alignment parameters in a dual-step EM algorithm.

Our sensitivity study uses randomly generated point-sets. We ensure that the point-sets have a clump structure by sampling the point positions from six partially overlapping Gaussian distributions with controlled variance. We have then added both new points at random positions, and, random point-jitter to the synthetic data. The randomly inserted points have been sampled from a uniform distribution. The positional jitter has been generated by displacing the points from their original positions by Gaussian measurement errors. The displacements have been randomly sampled from a circularly symmetric Gaussian distribution of zero mean and controlled standard deviation.

In Figure 1 we show the effect of increasing the number of randomly added points. In this experiment, we commence with a point-set of size 100. The plot shows the fraction of points correctly matched as a function of the number of randomly added points. The long-dashed curve, i.e. the one with gives the consistently lowest results, is the result of applying the Shapiro and Brady algorithm. Here the fraction of correct correspondences falls below 25% once the fraction of added clutter exceeds 2%. The dual-step EM used in our previous work which finds correspondences by explicitly aligning the points, is shown as a dot-dashed curve and performs best of all when the level of clutter is less than 20%. The remaining two curves show the results obtained with the two variants of our hierarchical correspondence algorithm detailed in Section 4.3. In the case of the dotted curve the cluster correspondences are computed using only the modal co-efficients of the between-cluster proximity matrix as described in Equation (12). The solid curve shows the results obtained if the eigenvalues are also used as described in Equation (13). There is little to distinguish the two methods. Both perform rather more poorly than the dual-step EM algorithm when the level of clutter is less than 20%. However, for larger clutter levels, they provide significantly better performance. The additional use of the eigenvlaues results in a slight improvement in performance.

Figure 2 investigates the effect of positional jitter. Here we plot the fraction of correct correspondence matches as a function of the standard deviation of the Gaussian position error added to the point-positions. We report the level of jitter using the ratio of the standard deviation of the Gaussian error distribution to the average closest inter-point distance. Here there is nothing to distinguish the behaviour of our hierarchical correspondence method from the dual-step alignment method. In each case the fraction of correct correspondences degrades slowly with increasing point-position jitter. However, even when the standard deviation of the position errors is 50% of the average minimum interpoint-distance then the fraction of correct correspondences is still greater than 50%. By contrast, the accuracy of the Shapiro and Brady method falls below 50% once the standard deviation of the positional error exceeds 10% of the minimum interpoint distance.

Our final set of experiments on synthetic data investigate the effect of diluting the cluster-structure of the point-sets. Here we have gradually moved the cluster-centres closer together and have investigated the effect on the fraction of correct correspondences when there is structural error present. The results are shown in figure 3. Here we represent the fraction of correct correspondences as a function of the overlap between the clusters. We have also included tests to show the performance of the algorithm when a 20% of clutter noise is added to the overlapping. The solid curve and the dashed curve respectively show the results obtained with the new method reported in this paper and the Shapiro and Brady method when the point-sets contain no clutter. The performance of the Shapiro and Brady method is poorer than the new method. Its sudden drop-off in performance is attributable to the effect of increased point-density as the clusters are overlapped. The dotted curve shows the result obtained with our new method when we match to the point-set with 20% clutter. Obviously the performance of the method is poorer than that obtained with the unclut-

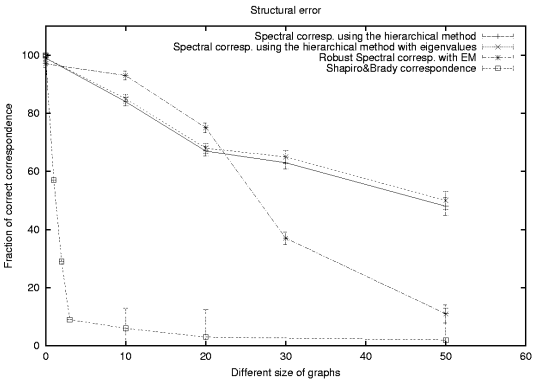

Fig. 1. Experimental results: structural error

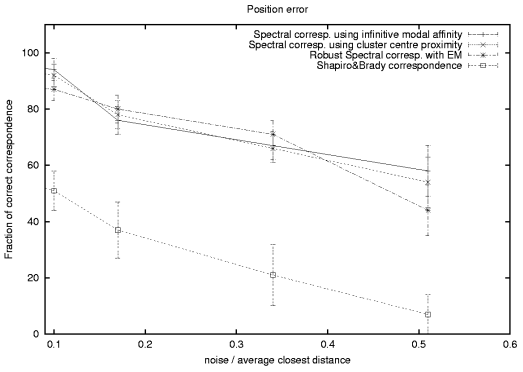

Fig. 2. Experimental results: position error

tered point-set. However, increased proximity of the clusters does not appear to significantly degrade performance.

5.2 Real World Data

In this section we provide some experiments with real world data. Our first experiment involves images from a gesture sequence of a hand. The images used in this study are shown in Figure 4. We commence by running the Canny edge detector over the images to locate the boundary of the hand. From this edge data, point features have been detected using the corner detector of Mokhtarian and Suomela [7]. The raw points returned by this method are distributed relatively uniformly along the outer edge of the hand and are hence not suitable for cluster analysis. We have therefore pruned the feature points using a curvature criterion. We have removed all points for which the curvature of the outline is smaller than a heuristically set threshold. Initially there are some 800 feature points, but after pruning this number is reduced to 271. The pruned feature-points are shown in

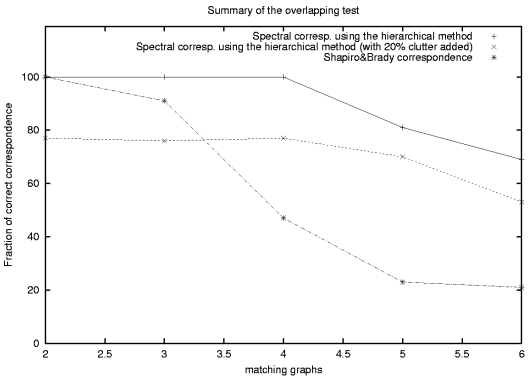

Fig. 3. Experimental results: final diagram for test on the cluster stability

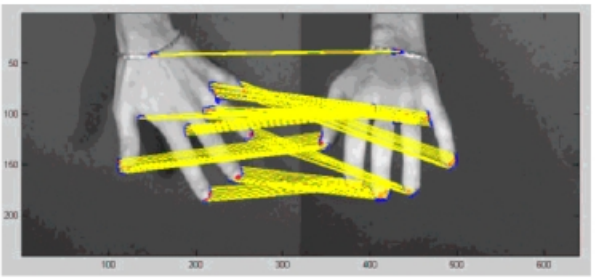

Fig. 4. Experimental results: real data experimentation

blue in the figure. They are clustered around the finger-tips and the points at which the fingers join the hand. After applying the clustering method, the set of centres shown in red is obtained. There are ten clusters in both images. The yellow lines between the two images show the detected correspondences. The fraction of correct correspondences is 81.2%.

Our second real world experiment involves a sequence of images obtained as a subject rotates and tilts their human head. The feature points here are highly non-planar. In Figure 5 we show the correspondences obtained. These are again good, and their appear to be no systematic problems.

A final example is shown in Figure 6 where we show the results obtained on an image pair from the roof-space of our lab. Here the correspondences are good despite the fact that there is no clear cluster-structure.

6 Conclusions

In this paper we have shown how constraints provided by the arrangement of modal groups of points can be used to improve the correspondence method of Shapiro and Brady [14]. The idea has been to use the modal co-efficients

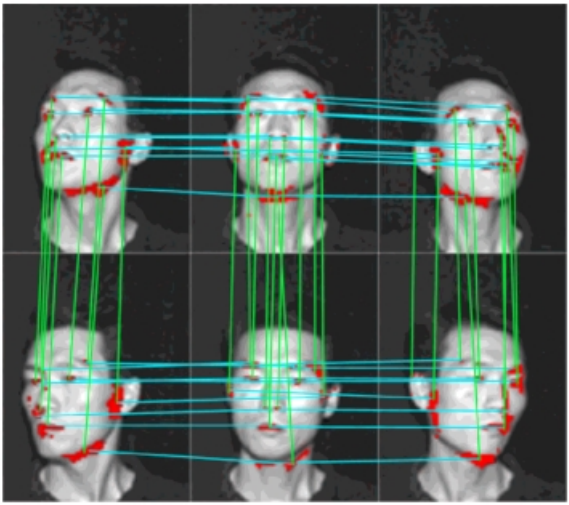

Fig. 5. Experimental results: real data experimentation

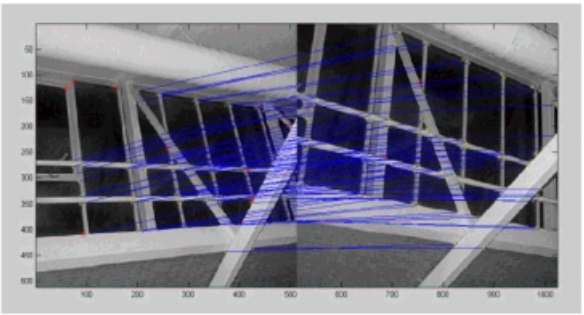

Fig. 6. Experimental results: real data experimentation

of the point-proximity matrix to establish the whereabouts of significant point groupings. We exploit these grouping to develop a hierarchical correspondence method. This is a two-step process. First, we use the spatial arrangements of the centre-points of the most significant groups to compute a between-cluster proximity matrix. The modal co-efficients of this between-cluster proximity matrix are used to establish correspondence probabilities between groups of points. Second, for each group of points we compute a within-cluster proximity matrix. The modal co-efficients of these within-cluster proxmity matrices are used to establish cluster-conditional point correspondence probabilities. Using the Bayes rule we combine these two sets of probabilities to compute individual point correspondence probabilities.

We have shown that while the Shapiro and Brady method fails once more than a few percent of clutter is added, the new method degrades more gracefully.

There are a number of ways in which the method described in this paper could be extended. One of the most important of these is to extend the method to line-pattern matching.

References

1. F.R.K. Chung, *Spectral Graph Theory*, CBMS series **92**, AMS Ed., 1997.
2. M.Carcassoni and E.R.Hancock , " Point Pattern Matching with Robust Spectral Correspondence ", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society Press*, **I**, pp. 649-655 , 2000.
3. A.P. Dempster, Laird N.M. and Rubin D.B., "Maximum-likelihood from incomplete data via the EM algorithm", *J. Royal Statistical Soc. Ser. B (methodological)*, **39**, pp. 1-38, 1977.
4. T. Hofmann and J.M. Buhmann, "Pairwise Data Clustering by Deterministic Annealing", PAMI(19), No. 1, January 1997, pp. 1-14.
5. R. Horaud and H. Sossa, "Polyhedral Object Recognition by Indexing", *Pattern Recognition*, **28**, pp. 1855-1870, 1995.
6. K.Inoue and K. Urahama, "Sequential fuzzy cluster extraction by a graph spectral method", *Pattern Recognition Letters*, **20**, pp. 699-705, 1999.
7. F. Mokhtarian and R. Suomela, "Robust Image Corner Detection Through Curvature Scale Space", *IEEE PAMI*, **20**:12, pp. 1376–1381, December 1998.
8. P Perona and W Freeman, "A Factorisation Approach to Grouping", ECCV 98, Vol 1, pp 655-670, 1998.
9. S. Sarkar and K.L. Boyer, K.L., "Quantitative Measures of Change Based on Feature Organization: Eigenvalues and Eigenvectors", CVIU(71), No. 1, July 1998, pp. 110-136.
10. S. Sarkar and P. Soundararajan, "Supervised Learning of Large Perceptual Organization: Graph Spectral Partitioning and Learning Automata", PAMI(22), No. 5, May 2000, pp. 504-525.
11. S. Sclaroff and A. Pentland, "Modal Matching for Correspondence and Recognition", *IEEE PAMI*, **17**:6, pp. 545–561, 1995.
12. G.L. Scott and H.C. Longuet-Higgins, "An algorithm for associating the features of 2 images", *Proceedings of the Royal Society of London Series B (Biological)*, **244**, pp. 21–26, 1991.
13. K. Sengupta and K.L.Boyer, "Modelbase partitioning using property matrix spectra", *Computer Vision and Image Understanding*, **70**:2, pp. 177-196, 1998.
14. L.S. Shapiro and J.M. Brady, "Feature-based correspondence - an eigenvector approach", *Image and Vision Computing*, **10**, pp. 283–288, 1992.
15. J. Shi and J.Malik, "Normalized cuts and image segmentation", *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 1997.
16. A. Shokoufandeh, S.J. Dickinson, K. Siddiqi and S.W. Zucker, "Indexing using a spectral encoding of topological structure", *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp.491-497, 1999.
17. S. Umeyama, "An eigen decomposition approach to weighted graph matching problems", *IEEE PAMI*, **10**, pp. 695–703, 1988.
18. Y. Weiss, "Segmentation using Eigenvectors: A Unifying View", International Conference on Computer Vision, pp. 975-982, 1999.