Spectral correspondence for point pattern matching

Marco Carcassoni*, Edwin R. Hancock

Department of Computer Science, University of York, York Y01 5DD, UK

Abstract

This paper investigates the correspondence matching of point-sets using spectral graph analysis. In particular, we are interested in the problem of how the modal analysis of point-sets can be rendered robust to contamination and drop-out. We make three contributions. First, we show how the modal structure of point-sets can be embedded within the framework of the EM algorithm. Second, we present several methods for computing the probabilities of point correspondences from the modes of the point proximity matrix. Third, we consider alternatives to the Gaussian proximity matrix. We evaluate the new method on both synthetic and real-world data. Here we show that the method can be used to compute useful correspondences even when the level of point contamination is as large as 50%. We also provide some examples on deformed point-set tracking. © 2002 Published by Elsevier Science Ltd on behalf of Pattern Recognition Society.

Keywords: Point pattern matching; Proximity matrix; Eigenvectors; Sequence analysis; Robust error kernel

1. Introduction

Spectral graph theory is a term applied to a family of techniques that aim to characterise the global structural properties of graphs using the eigenvalues and eigenvectors of either the adjacency matrix or the closely related Laplacian matrix [1]. Although the subject has found widespread use in a number of areas including structural chemistry and routing theory, there have been relatively few applications in the computer vision literature. The reason for this is that although elegant, spectral graph representations are notoriously susceptible to the effect of structural error. In other words, spectral graph theory can furnish very efficient methods for characterising exact relational structures, but soon breaks down when there are spurious nodes and edges in the graphs under study.

Matrix factorisation methods have been extensively used to solve problems posed as least squares estimation. These include a host of techniques for point-set matching [2–6], stereo [7] and structure-from-motion [8]. However, these methods operate with co-ordinate position and are not concerned with the relational information conveyed by proximity matrices or adjacency graphs. Spectral methods for graph analysis invariably commence by computing the Laplacian matrix. This is closely related to the node adjacency matrix. The diagonal elements of the Laplacian matrix are equal to the degree of the nodes (vertices) and the off diagonal elements are unity if the corresponding nodes are connected by an edge, and are zero otherwise. However, it is also common to work with proximity or property matrices where the off diagonal elements reflect the difference in node attributes such as position [6] or orientation [9]. Once a matrix characterisation of the graph is to hand then the eigenvalues and eigenvectors are computed. The main idea behind spectral graph theory is to use the distribution of eigenvalues to provide a compact summary of graph-structure.

In the computer vision literature there have been a number of attempts to use spectral properties for graph-matching, object recognition and image segmentation. Umeyama [2] has developed a singular value decomposition method for finding the least-squares permutation matrix between the adjacency matrices of two graphs. This eigendecomposition method can be used for exact graph-matching. That is to say, it finds correspondences between the nodes of graphs

---

* Corresponding author. Tel.: +44-1904-432795; fax: +44-1904-432767.

E-mail address: marco@cs.york.ac.uk (M. Carcassoni).

of the same size and edge-structure [2]. Borrowing ideas from structural chemistry, Scott and Longuet-Higgins were among the first to use spectral methods for correspondence analysis [5]. They showed how to recover correspondences via singular value decomposition on the point association matrix between different images. The method can work with point-sets of different size, but is sensitive to the degree of rotation, and to a lesser extent the amount of scaling, between the two point-sets. In order to overcome these problems with the Scott and Longuet-Higgins method, Shapiro and Brady [6] have developed a method which uses the intra-image point proximity matrix. The method is more closely in keeping with the spirit of spectral graph theory. Brady and Shapiro's method commences by computing a point-proximity matrix by using a Gaussian function of the distance between points in the same image. A modal representation is constructed by finding the eigenvalues and eigenvectors of the proximity matrix. Correspondences are located by comparing the ordered eigenvectors of the proximity matrices for different images. The eigenvectors of the proximity matrices can be viewed as the basis vectors of an orthogonal transformation on the original point identities. In other words, the components of the eigenvectors represent the affinities of the original points to the eigenmodes of the proximity matrix. Provided that the point-sets are of the same size, then the correspondences delivered by the Shapiro and Brady method are relatively robust to random point jitter and to affine rotations and scalings. The work of Scott and Longuet-Higgins [5], and, Shapiro and Brady [6] can be viewed as providing the foundations on which the more recent modal matching methods of Sclaroff and Pentland [4] and Cootes et al. [3] build. Becuase they rely on singular value decomposition the methods are also intimately related to Procrustes shape analysis [10].

Since it uses point proximity information, Shapiro and Brady's method can be viewed as operating in the attribute domain rather than the symbolic or structural domain. By contrast, Horaud and Sossa [11] have adopted a purely structural approach to the recognition of line-drawings. Their representation is based on the immanantal 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. In another application involving indexing into large data-bases, Sengupta and Boyer [9] have used property matrix spectra to characterise line-patterns. Various attribute representations are suggested and compared. Shokoufandeh et al. [12] have shown how graphs can be encoded using local topological spectra for shape recognition from large data-bases. Sarkar and Soundararajan [13] have shown how graph-spectral methods can be combined with automata to learn grouping structure. Finally, a number of authors have used spectral methods to perform pairwise clustering on image data. Shi and Malik [14] use the second eigenvalue to segment images by performing an eigendecomposition on a matrix of pairwise attribute differences using the iterative normalised cut method.

Sarkar and Boyer [15] and Freeman and Perona [16] have developed matrix factorisation methods for line-grouping. In a recent paper, Weiss [17] 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 [18] 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 [19].

The focus of this paper is the use of property matrix spectra for correspondence matching. As mentioned above, spectral methods offer an attractive route to correspondence matching since they provide a representation that can be used to characterise graph structure at the global level. If used effectively, the spectral representation can be used for rapid matching by comparing patterns of eigenvalues or eigenvectors. Conventional graph-matching methods, on the other hand, rely on local structural decompositions [20–24]. Correspondence analysis is achieved through the iterative propagation of local consistency constraints with the hope of achieving global consistency Moreover, accuracy is traded against the efficiency gains achieved by using increasingly localised structures. Invariably global consistency is only assessed using the edges of the graphs. However, the increased fidelity of representation achieved using a global spectral representation must be weighed against their relative fragility to the addition of noise and clutter. For instance, although the methods of Horaud and Sossa [11] and Shapiro and Brady [6] work well for graphs that are free of structural contamination, they do not work well when the graphs are of a different size. Moreover, the method of Sengupta and Boyer [9], although relatively robust is not concerned with detailed correspondence analysis.

Our aim in this paper is to consider how spectral methods can be rendered robust for correspondence matching with point-sets which contain significant noise and contamination. To do this we make three contributions. First we cast the problem of correspondence matching in the framework of the EM algorithm. This gives us scope for outlier reject via weighting in the expectation step. The framework used as the starting point for our study is the dual-step EM algorithm of Cross and Hancock [25]. This work has shown how local relational constraints can be embedded in an EM algorithm for point alignment under affine and perspective distortion. We commence by showing how graph-spectra can be used to compute the required correspondence probabilities in a manner which is both efficient and global. Our second contribution is to consider how to make the initial computation of the point proximity matrix more robust to contamination by the addition of clutter and the relative movement of points due to measurement noise. Finally, we consider how to compare the node mixing angles in a robust manner. Here we model the effects of confused or-

dering of the eigenvalues and noisy estimates of the mixing eigenvectors.

The outline of this paper is as follows. In Section 2 we define some of the formal ingredients of our method. Section 3 details the dual-step EM algorithm used to align point-sets under affine geometry. In Section 4 we provide details of how to use the modal structure of the proximity matrix to compute the correspondence probabilities required by the dual-step EM algorithm. Section 5 provides a sensitivity study for the new method on synthetic point-sets. This study focuses on how the best choice of weighting function for computing the proximity matrix, and how to use the modal structure to compute correspondence probabilities. In Section 6 we demonstrate the utility of the new alignment method on real-world image sequences. Finally, Section 7 offers some conclusions and suggests some future lines of investigation.

2. Prerequisites

The aim in this paper is to use the dual-step EM algorithm of Cross and Hancock [25] to render the process of spectral pattern matching robust. Before we detail the algorithm, we provide some of the formal ingredients of the method.

2.1. Affine geometry

Suppose that $\Phi^{(n)}$ is the geometric transformation that best aligns a set of image feature points **w** with their counterparts in a model. Each point in the image data set is represented by an augmented position vector $\vec{w}_i = (x_i, y_i, 1)^{\mathrm{T}}$ where i is the point index. This augmented vector represents the two-dimensional point position in a homogeneous co-ordinate system. We will deal only with point-sets which lie on a single plane in the image. In the interests of brevity we will denote the entire set of image points by $\mathbf{w} = \{\vec{w}_i, \forall i \in \mathscr{D}\}$ where $\mathscr{D}$ is the point set. The corresponding fiducial points constituting the model are similarly represented by $\mathbf{z} = \{\vec{z}_j, \forall j \in \mathscr{M}\}$ where $\mathscr{M}$ denotes the index-set for the model feature-points $\vec{z}_j$.

In this paper we confine our attention to affine transformations. The affine transformation has six free parameters. These model the two components of translation of the origin on the image plane, the overall rotation of the co-ordinate system, the overall scale together with the two parameters of shear. These parameters can be combined succinctly into an augmented matrix that takes the form

$$\Phi^{(n)} = \begin{pmatrix} \phi_{1,1}^{(n)} & \phi_{1,2}^{(n)} & \phi_{1,3}^{(n)} \\ \phi_{2,1}^{(n)} & \phi_{2,2}^{(n)} & \phi_{2,3}^{(n)} \\ 0 & 0 & 1 \end{pmatrix}. \tag{1}$$

With this representation, the affine transformation of co-ordinates is computed using the following matrix multiplication:

$$\vec{w}_i^{(n)} = \Phi^{(n)} \vec{w}_i. \tag{2}$$

Clearly, the result of this multiplication gives us a vector of the form $\vec{w}_i^{(n)} = (x, y, 1)^{\mathrm{T}}$. The superscript n indicates that the parameters are taken from the nth iteration of our algorithm.

2.2. Correspondences

The recovery of the parameters of the transformation matrix Φ, requires correspondences between the point-sets. In other words, we need to know which point in the data aligns with which point in the model. This set of correspondences between the two point sets is denoted by the function $f^{(n)} : \mathscr{D} \rightarrow \mathscr{M}$ from the nodes of the data-graph to those of the model graph. According to this notation the statement $f^{(n)}(i) = j$ indicates that there is a match between the node $i \in \mathscr{D}$ of the model-graph to the node $j \in \mathscr{M}$ of the data-graph at iteration n of the algorithm.

3. The dual-step EM algorithm

Cross and Hancock's contribution was to present an extension of the standard EM algorithm in which the structural consistency of correspondences matches can be used to gate contributions to the expected log-likelihood function [25]. This idea is closely related to the hierarchical mixture of experts algorithm of Jordan and Jacobs [26]. However, the method uses a dictionary method for computing the correspondence probabilities which is both localised and time consuming. The aim here is to replace the dictionary-based method used to compute the probabilities with a robust spectral method.

3.1. Expected log-likelihood

Following, Cross and Hancock we seek both correspondence matches (i.e. the function f) and transformation parameters (i.e the matrix Φ) which maximise the expected log-likelihood

$$Q(\Phi^{(n+1)}|\Phi^{(n)}) = \sum_{i \in \mathscr{D}} \sum_{j \in \mathscr{M}} P(\vec{z}_j | \vec{w}_i, \Phi^{(n)}) \varsigma_{i,j}^{(n)}$$

$$\times \ln p(\vec{w}_i | \vec{z}_j, \Phi^{(n+1)}). \tag{3}$$

The meaning of this expected log-likelihood function requires further comment. The measurement densities $p(\vec{w}_i | \vec{z}_j, \Phi^{(n+1)})$ model the distribution of error-residuals between the data-point position $\vec{w}_i$ and the model point position $\vec{z}_j$ at iteration n of the algorithm. The log-likelihood contributions at iteration $n + 1$ are weighted by the a posteriori measurement probabilities $P(\vec{z}_j | \vec{w}_i, \Phi^{(n)})$ computed at the previous iteration n. The individual contributions to the expected log-likelihood

function are gated by the structural matching probabilities $\zeta_{i,j}^{(n)}$. These probabilities measure the consistency of the pattern of correspondences when the match $f^{(n)}(i) = j$ is made. Their modeling is the topic of Section 4.

3.2. Expectation

In the expectation step of the EM algorithm the a posteriori probabilities of the missing data (i.e. the model-graph measurement vectors, $\vec{z}_j$) are updated by substituting the updated point position vectors into the conditional measurement distribution. Using the Bayes rule, we can re-write the a posteriori measurement probabilities in terms of the components of the corresponding conditional measurement densities:

$$P(\vec{z}_j|\vec{w}_i, \Phi^{(n)}) = \frac{\alpha_j^{(n)} p(\vec{w}_i|\vec{z}_j, \Phi^{(n)})}{\sum_{j' \in \mathcal{M}} \alpha_{j'}^{(n)} p(\vec{w}_i|\vec{z}_{j'}, \Phi^{(n)})}. \qquad (4)$$

The mixing proportions are computed by averaging the a posteriori probabilities over the set of data-points, i.e.

$$\alpha_j^{(n+1)} = \frac{1}{|\mathscr{D}|} \sum_{i \in \mathscr{D}} P(\vec{z}_j|\vec{w}_i, \Phi^{(n)}).$$

In order to proceed with the development of a point registration process we require a model for the conditional measurement densities, i.e. $p(\vec{w}_i|\vec{z}_j, \Phi^{(n)})$. Here we assume that the required model can be specified in terms of a multivariate Gaussian distribution. The random variables appearing in these distributions are the error residuals for the position predictions of the jth model point delivered by the current estimated data-point positions. Accordingly we write

$$p(\vec{w}_i|\vec{z}_j, \Phi^{(n)}) = \frac{1}{2\pi\sqrt{|\Sigma|}}$$
$$\times \exp\left[-\frac{1}{2}(\vec{z}_j - \Phi^{(n)}\vec{w}_i)^{\mathrm{T}} \Sigma^{-1}(\vec{z}_j - \Phi^{(n)}\vec{w}_i)\right].$$
$$(5)$$

In the above expression Σ is the variance–covariance matrix for the position errors.

3.3. Maximisation

The dual step EM algorithm iterates between the two interleaved maximisation steps for alignment parameter estimation and estimating correspondence assignments.

3.3.1. Maximum a posteriori probability matches

Point correspondences are sought so as to maximise the a posteriori probability of structural match. The update formula is

$$f^{(n+1)}(i) = \arg\max_{j \in \mathcal{M}} P(\vec{z}_j|\vec{w}_i, \Phi^{(n)})\zeta_{i,j}^{(n)}. \qquad (6)$$

3.3.2. Maximum likelihood alignment

In the case of affine geometry, the transformation is linear in the parameters. This allows us to locate the maximum-likelihood parameters directly by solving the following system of saddle-point equations for the independent affine parameters $\phi_{k,l}^{(n+1)}$ running over the indices $k = 1, 2$ and $l = 1, 2, 3$:

$$\frac{\partial Q(\Phi^{(n+1)}|\Phi^{(n)})}{\partial \phi_{k,l}^{(n+1)}} = 0. \qquad (7)$$

For the affine transformation the set of saddle-point equations is linear, and are hence easily solved by using matrix inversion. The updated solution matrix is given by

$$\Phi^{(n+1)} = \left[\sum_{i \in \mathscr{D}} \sum_{j \in \mathcal{M}} P(\vec{z}_j|\vec{w}_i, \Phi^{(n)})\zeta_{i,j}^{(n)}\vec{w}_i U^{\mathrm{T}} \vec{w}_i^{\mathrm{T}} \Sigma^{-1}\right]^{-1}$$
$$\times \left[\sum_{i \in \mathscr{D}} \sum_{j \in \mathcal{M}} P(\vec{z}_j|\vec{w}_i, \Phi^{(n)})\zeta_{i,j}^{(n)}\vec{z}_j U^{\mathrm{T}} \vec{w}_i^{\mathrm{T}} \Sigma^{-1}\right], \quad (8)$$

where the elements of the matrix U are the partial derivatives of the affine transformation matrix with respect to the individual parameters, i.e.

$$U = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \qquad (9)$$

This allows us to recover a set of improved transformation parameters at iteration $n + 1$. Once these are computed, the a posteriori measurement probabilities may be updated by applying the Bayes formula to the measurement density function. The update procedure involves substituting the parameter matrix of Eq. (1) into the Gaussian density of Eq. (6) and applying the Bayes theorem.

4. Spectral methods for correspondence matching

The aim in this section is to show how the dual step EM algorithm can be used to render the process of computing correspondences between the nodes of the graphs robust to structural noise and measurement error. The spectral approach to correspondence commences by enumerating a point proximity matrix. This is a continuous 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. Once the proximity matrix is to hand, then correspondences are located by computing its eigenvectors. The components of the eigenvectors or columns of the modal matrix represent the affinities of the different points to the individual modes of the point

proximity matrix. The co-efficients appearing in the rows of the modal matrix represent the affinities of individual points to the different eigenmodes. Hence, we can locate point correspondences by searching for rows of the transformation matrix which have maximal similarity.

Unfortunately there are two drawbacks with the spectral method of correspondence. Firstly, there is no clear reason to use Gaussian weighting in favour of possible alternatives. Moreover, the 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.

The aim in this section is to address these two problems. We commence by considering alternatives to Gaussian weighting. Next we suggest how the comparison of the eigenvectors can be effected in a manner which is robust to measurement error and poor ordering of the eigenvalues. The contribution is therefore to show how the correspondence probabilities $\zeta_{i,j}^{(n)}$ can be computed in an efficient and robust manner.

4.1. Point proximity matrix

In this section, we suggest various alternative ways of constructing the weighted point-proximity matrix. The role of the weighting function is to model the probability of adjacency relations between points. In Shapiro and Brady's original work the weighting function was the Gaussian [6]. Here we consider various alternative weighting functions suggested by the robust statistics literature.

According to robust statistics, the effects of outliers can be controlled by weighting according to the error-residual. Suppose that $\Gamma_s(\eta)$ is a weighting function defined on the error-residual η. The parameter s controls the width of the weighting kernel. Associated with the weighting function is an error-kernel which is defined to be

$$\rho_s(\eta) = \int_{-\infty}^{\eta} \eta' \Gamma_s(\eta') \, d\eta'. \tag{10}$$

There are many choice of possible weighting functions described in the literature. However, they can be classified according to a broad-based taxonomy based on the derivative $\rho_s'(\eta)$ of the error-kernel. If the derivative is monotonically increasing, then the weighting function is said to be increasing. If the derivative is asymptotically constant, then the weighting function is said to be sigmoidal. Finally, if the derivative asymptotically approaches zero then the weighting function is said to be re-descending.

In this section, we investigate several weighting functions which fall into these different classes. Fig. 1 shows the graphs of the functions.

4.1.1. Gaussian weighted proximity matrix
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 at iteration n of the algorithm is given by

$$H_D^{(n)}(i, i') = \exp\left[-\frac{1}{2s^2} \| w_i^{(n)} - w_{i'}^{(n)} \|^2\right]. \tag{11}$$

This weighting function is re-descending.

4.1.2. Sigmoidal proximity matrix
To provide an example of a sigmoidal weighting function, we consider the proximity matrix generated by the hyperbolic tangent function

$$H_D^{(n)}(i, i') = \frac{2}{\pi \| w_i^{(n)} - w_{i'}^{(n)} \|} \log \cosh\left[\frac{\pi}{s} \| w_i^{(n)} - w_{i'}^{(n)} \|\right]. \tag{12}$$

4.1.3. Increasing weighting function
Here we use the increasing weighting function

$$H_D^{(n)}(i, i') = \left[1 + \frac{1}{s} \| w_i^{(n)} - w_{i'}^{(n)} \|\right]^{-1} \tag{13}$$

to generate the proximity matrix.

4.1.4. Euclidean weighting function
Here we investigate the effect of using a weighting function that decreases linearly with distance. We use the following piecewise specification to define a trapezoid function $H_D^{(n)}(i, i') = T(\| w_i^{(n)} - w_{i'}^{(n)} \|)$ where

$$T(\eta) = \begin{cases} 1 & \text{if } \eta < s_1, \\ 1 - \frac{1}{s_2 - s_1}[\eta - s1] & \text{if } s_1 < \eta < s_2, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

Here s_1 is the half-width of the ceiling of the function and s_2 is the half-width of the base.

4.2. Correspondences

The modal structure of the two point-sets is found by solving the eigenvalue equation

$$|H - \lambda I| = 0$$

and computing the associated eigenvectors using the equation

$$HE_l = \lambda_l E_l,$$

where λ_l is the lth eigenvalue of the matrix H and E_l is the corresponding eigenvector. We order the vectors according to the size of the associated eigenvalues. The ordered column-vectors are used to construct a modal matrix
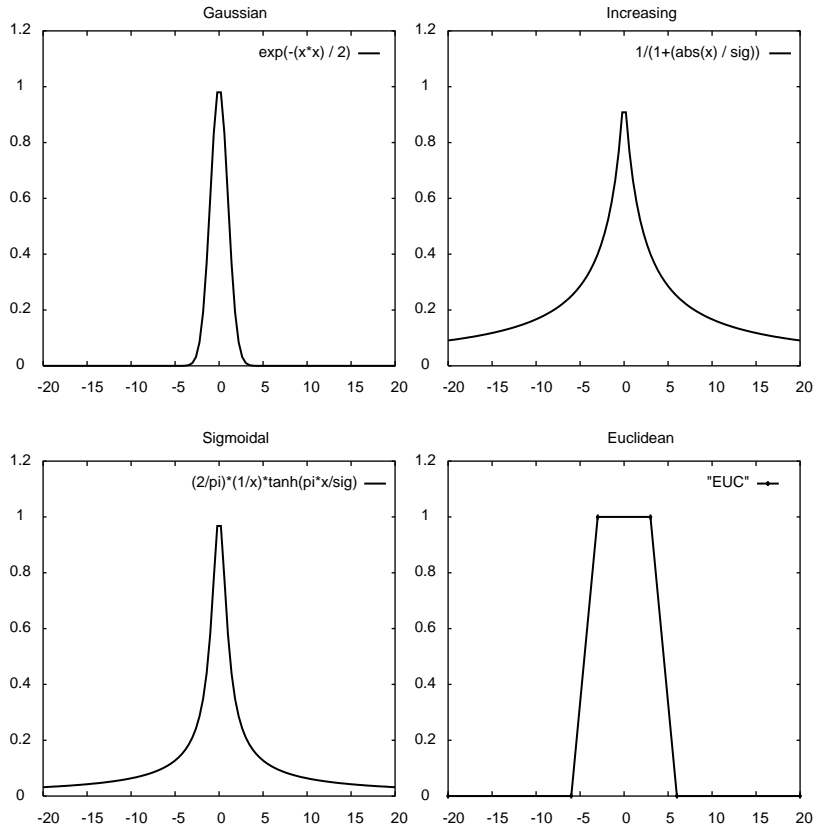
$$V = (E_1, E_2, E_3, \ldots).$$

Fig. 1. Weighting functions.

The column index of this matrix refers to the order of the eigenvalues while the row-index is the index of the original point-set. This modal decomposition is repeated for both the model and transformed data point-sets to give a data-point modal matrix

$$V_D^{(n)} = (E_1^D, E_2^D, E_3^D, \ldots, E_{|\mathscr{D}|}^D)$$

and a model-point modal matrix

$$V_M = (E_1^M, E_2^M, E_3^M, \ldots, E_{|\mathscr{M}|}^M).$$

Since the two point-sets are potentially of different size, we truncate the modes of the larger point-set. This corresponds to removing the last $\|\mathscr{D}\| - |\mathscr{M}\|$ rows and columns of the larger matrix. The resulting matrix has $o = \min[\mathscr{D}, \mathscr{M}]$ 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 eigenmodes.

4.2.1. Shapiro and Brady

Based on this eigendecomposition Shapiro and Brady [6] find correspondences by comparing the rows of the model matrices V_M and V_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:

$$\zeta_{i,j}^{(n)} = \begin{cases} 1 & \text{if } j = \arg\min_{j'} \sum_{l=1}^{o} \|V_D^{(n)}(i,l) - V_M(j',l)\|^2 \\ 0 & \text{otherwise.} \end{cases}$$

(15)

4.2.2. Correspondence probabilities

Rather than using the binary correspondence pattern of Shapiro and Brady to gate contributions to the expected log-likelihood function, we can use the elements of the two modal matrices to compute the probabilities of correspondence match. This is done by comparing the elements of the two matrices on a row-by-row basis. A simple way of

computing the probabilities is to assume that the vectors are subject to Gaussian measurement errors. As a result, the probability that data-point i is in correspondence with model data-point j is equal to

$$\zeta_{i,j}^{(n)} = \frac{\exp[-\mu \sum_{l=1}^{o} \|V_D^{(n)}(i,l) - V_M(j,l)\|^2]}{\sum_{j' \in \mathcal{M}} \exp[-\mu \sum_{l=1}^{o} \|V_D^{(n)}(i,l) - V_M(j',l)\|^2]}. \quad (16)$$

4.2.3. Robust correspondence probabilities

The shortcoming of this method for computing the correspondence probabilities is the effect of outlier measurement errors. When there is a significant difference between one or more of the components of the eigenvectors, then these errors dominate the argument of the exponentials. This will have the tendency to flatten the distribution and will result in ambiguity and equivocation concerning the pattern of correspondences. 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 we define the correspondence probability to be

$$\zeta_{i,j}^{(n)} = \frac{\sum_{l=1}^{o} \exp[-\mu \|V_D^{(n)}(i,l) - V_M(j,l)\|^2]}{\sum_{j' \in \mathcal{M}} \sum_{l=1}^{o} \exp[-\mu \|V_D^{(n)}(i,l) - V_M(j',l)\|^2]}. \quad (17)$$

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

4.2.4. Confused ordering of the eigenvalues

The method described above compares the components of the eigenvectors on a vector-by-vector basis. This assumes that there is no error in the computation of eigenvalues, and hence no leakage between the components of the vectors. Recall that the magnitudes of the eigenvalues establish an order for the components of the eigenvectors in the modal matrix. We model errors in the placement of the components of the eigenvectors by assuming that there is an error associated with the eigenvalues. This error distribution is used as a prior on the order of the components of the eigenvectors.

To commence let $P(f(i) = j | \omega_{\lambda_D} = \omega_{\lambda_M})$ represent the probability that the data point i is in correspondence with the model point j, given that the data-graph modal matrix row indexed ω_{λ_D} matches against the model-graph modal matrix row indexed ω_{λ_M}. This is the confusion probability for the spectral correspondence process. Further, let $P(\omega_{\lambda_D} = \omega_{\lambda_M})$ denote the prior probability that the data-graph modal matrix row indexed ω_{λ_D} matches against the model-graph modal matrix row indexed ω_{λ_M}. Our viewpoint is that the ordering of the eigenvalues is uncertain, and in consequence we do not know how to associate rows in the two matrices. Accordingly, we use the Bayes formula to expand the point correspondences over the missing eigenvalues correspondence.

As a result we write

$$\zeta_{i,j}^{(n)} = \sum_{\omega_{\lambda_D}} \sum_{\omega_{\lambda_M}} P(\omega_{\lambda_D} = \omega_{\lambda_M}) P(f(i) = j | \omega_{\lambda_D} = \omega_{\lambda_M}). \quad (18)$$

To develop this idea further, we require models of the confusion probabilities and the prior for the ordering of the eigenvalues. We assume the following Gaussian model for the deviation between the eigenvalues:

$$P(\omega_{\lambda_D} = \omega_{\lambda_M}) = \frac{1}{\sqrt{2\pi}\sigma_\lambda} \exp\left[-\frac{1}{2\sigma_\lambda^2}(\lambda_D - \lambda_M)^2\right]. \quad (19)$$

The confusion probabilities are assumed to follow the distribution outlined in the previous subsection. As a result we write

$$\zeta_{i,j}^{(n)} =$$

$$\frac{\sum_{\omega_{\lambda_D}=1}^{|\mathcal{D}|} \sum_{\omega_{\lambda_M}=1}^{|\mathcal{M}|} \exp[-1/2\sigma_\lambda^2(\lambda_D-\lambda_M)^2] S_{i,j}^{\omega_{\lambda_D},\omega_{\lambda_D}}}{\sum_{j' \in \mathcal{M}} \sum_{\omega_{\lambda_D}=1}^{|\mathcal{D}|} \sum_{\omega_{\lambda_M}=1}^{|\mathcal{M}|} \exp[-1/2\sigma_\lambda^2(\lambda_D-\lambda_M)^2] S_{i,j'}^{\omega_{\lambda_D},\omega_{\lambda_D}}}, \quad (20)$$

where

$$S_{i,j}^{\omega_{\lambda_D},\omega_{\lambda_D}} = \exp[-\mu\|V_D^{(n)}(i,\omega_{\lambda_D}) - V_M(j,\omega_{\lambda_M})\|^2]. \quad (21)$$

5. Experiments

In this section, we investigate the performance of the different methods of spectral correspondence reported in this paper. There are three aspects to this experimental study. First we consider the effect of varying the way in which the elements of the point proximity matrix are calculated. Next, we compare the four methods for computing correspondences from the eigenvectors of the proximity matrix. Finally, we show the benefits to be gained from embedding the correspondence process in the EM algorithm.

Our experiments are conducted with random point sets. We investigate two sources of error. The first of these is random measurement error or point-position jitter. Here we subject the positions of the points to Gaussian measurement error. The parameter of the noise process is the standard deviation of the point position error. The second source of error is structural. Here we remove random points. This is the most destructive type of error for spectral methods.

5.1. Point proximity matrix

Here we investigate the effect of varying the point proximity matrix. The aim is to determine which of the weighting functions returns correspondences which are the most robust to point-position jitter. Fig. 2 shows the fraction of correct correspondences as a function of the standard deviation of the added Gaussian position errors. The
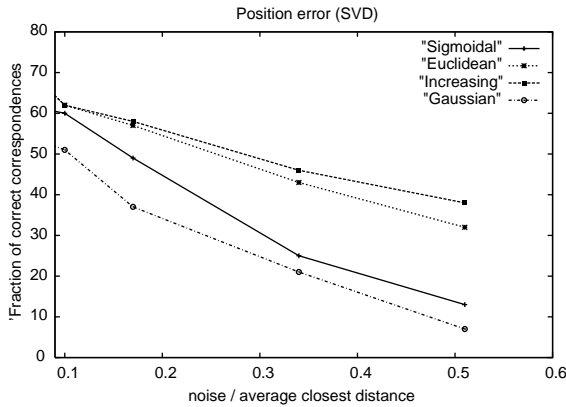
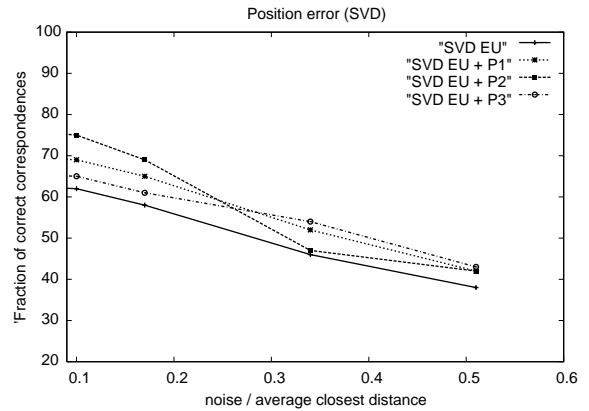Fig. 2. Effect of weighting function on correspondence error.

Fig. 3. Effect of matching probabilities on correspondence error.

standard deviation is recorded as a fraction of the average closest point distance. The best performance is obtained with the increasing weighting function and the trapezoidal Euclidean weighting function. The next best method is the sigmoidal weighting function. The poorest performance is returned by the Gaussian proximity matrix. However, the two best performing weighting functions give a margin of improvement of some 20% over the Gaussian weighting function.

5.2. Correspondence probabilities

In this section, we compare the results obtained using different methods for estimating correspondences from the eigenvectors of the point-proximity matrix. In each case we use the increasing weighting function to construct the proximity matrix. We use only the SVD methods to find correspondences. The correspondences are selected on the basis of maximum probability using the different schemes outlined in Section 4.2. The effect of embedding the correspondence probabilities in the EM algorithm is deferred until Section 5.3.

Here we investigate the effect of point position jitter. Fig. 3 shows the fraction of correct correspondences as a function of the standard deviation of the Gaussian measurement error. The results are reported using the ratio of the standard deviation to the average closest inter-point distance. The four curves correspond to the standard method of Shapiro and Brady and the three probabilistic methods outlined in Section 4.1. All four methods degrade as the noise standard deviation increases. The three probabilistic methods offer significant improvement over the method of Shapiro and Brady. Although there is little to distinguish the overall performance of the three probabilistic methods, the following pattern emerges: The robust probabilities (Section 4.2.3) work best for low noise levels; the eigenvalue confusion model (Section

4.3.3) works best at high noise levels; the simplest probabilistic method (Section 4.3.1) gives the best overall performance.

5.3. EM algorithm

In this section, we turn our attention to the effect of embedding the correspondence probabilities in the EM algorithm. We compare three different algorithms. The first of these is the standard SVD method of Shapiro and Brady. The second is the alignment of the point-sets using the standard EM algorithm. This algorithm does not use any information concerning the pattern of correspondences. Contributions to the log-likelihood function are weighted only by the a posteriori alignment probabilities. In other words, we set $\zeta_{i,j}^{(n)} = 1$ for all values of i and j. Finally, we consider the new approach to spectral correspondence developed in this paper. We refer to these algorithms as SVD, EM and EM+SVD, respectively.

We commence in Fig. 4 by showing the effect of positional jitter on the fraction of correct correspondences. The plot shows the fraction of correct matches as a function of the size of the point sets. Here the standard deviation of the Gaussian noise is kept fixed and the size of the point-sets is increased. The two point sets are of the same size. There is no addition or deletion of points. The size of the point-sets is increased by adding new points at random positions. As a result the point-density increases. The main effects to note are the following. Firstly, all three methods degrade with the increasing size of the point-sets. One of the reasons for this is that as the size of the point-sets is increased, then so the average inter-point distance decreases. In other words, the degree of relative displacement increases with the increasing number of points. The second feature to note is that the new method (i.e. SVD+EM) consistently gives the best results. Moreover, the performance of the pure SVD method degrades rapidly as the
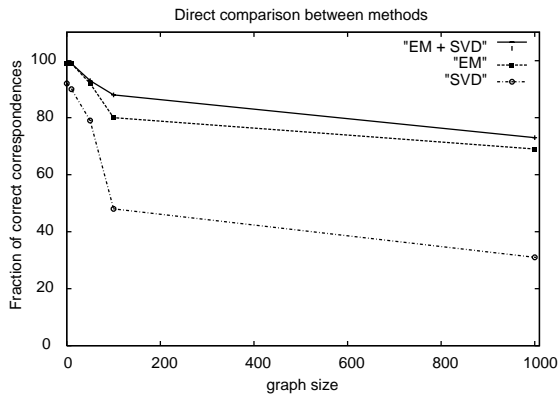
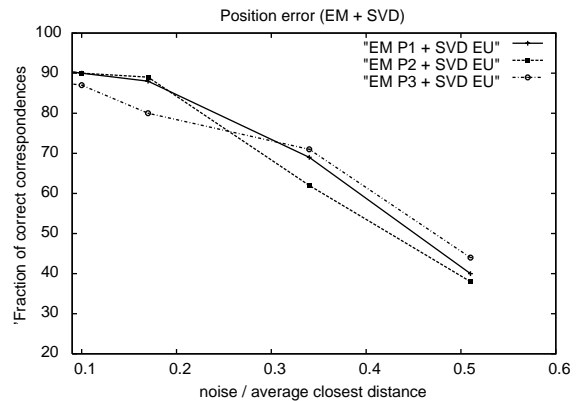Fig. 4. Effect of graph-size on correspondence error.

Fig. 6. Effect of point position error on correspondence error.
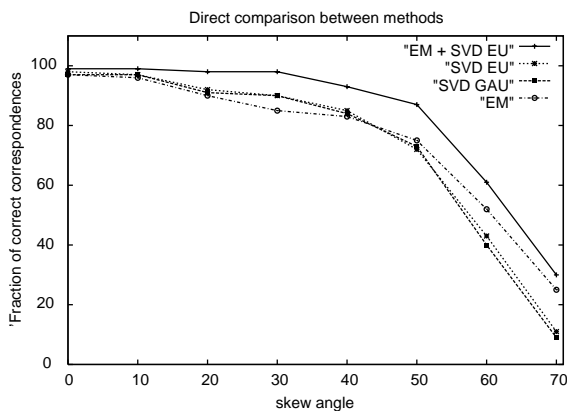

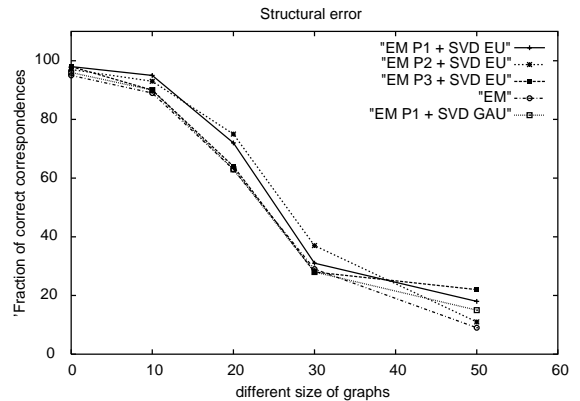Fig. 5. Effect of skew angle on correspondence error.

Fig. 7. Effect of clutter on correspondence error.

relative displacement of the points is increased. Finally, the SVD+EM method outperforms the EM method by a useful margin.

Next we investigate the effect of controlled affine skew of the point-sets. Fig. 5 shows the fraction of correct correspondences as a function of the skew angle in degrees. Here we compare the standard EM method, the EM method with a trapezoidal weighting function, and the use of pure SVD with a Gaussian weighting function and a trapezoidal weighting function. Here the embedding of the new proximity matrix in the EM algorithm offers clear advantages. There is a 10% margin of improvement for all skew angles.

We now turn to the effect of positional jitter. Fig. 6 compares the effect of using the three different ways of computing the correspondence probabilities. There is no obvious pattern of improvement. However, comparing with Fig. 3 it is clear that the EM method gives a mar-

gin of improvement of some 20% over the use of SVD alone.

Finally, we investigate the effects of point-set contamination. Here we add increasing fractions of cluster nodes to the point-sets. We commence by noting that the pure SVD method fails as soon as clutter is added. Hence we cannot plot any sensitivity data from the method. Fig. 7 shows the effect of adding controlled fractions of clutter to point-sets of increasing size. The main conclusion to note is that SVD+EM outperforms EM by about 5–10%. Also, as the fraction of clutter increases then so the performance degrades.

6. Real world data

Our final piece of experimental work focuses on real-world data. Here we have matched images from a gesture sequence of a moving hand. The feature points
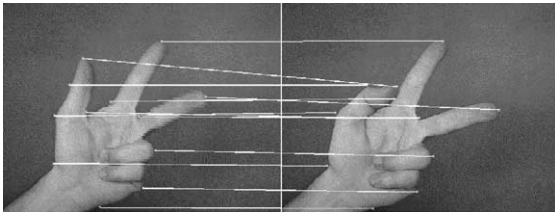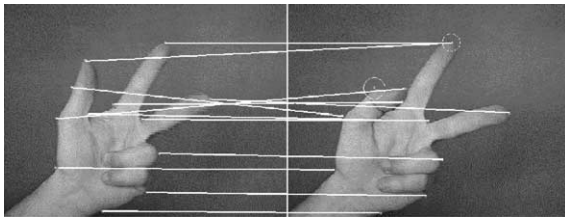
Fig. 8. Final matches-EM.

Fig. 9. Final matches-SVD.

Fig. 10. Matches—dual-step EM+SVD.

in these images are points of maximum curvature on the outline of the hand. Fig. 8 shows the final configuration of correspondence matches obtained using our new method. For comparison, Fig. 9 shows the correspondence matches obtained with SVD. In the case of the EM method, all the correspondence matches are correct, while in the case of SVD the two correspondences marked with circles are in error.

To illustrate the effectiveness of the method for feature, we have extended these experiments to a longer motion sequence in which a hand is clenched to form a fist. Figs. 10 and 11, respectively, show the sequence of correspondence matches obtained with the dual-step EM algorithm and the SVD methods of Shapiro and Brady. In each panel the image on the left-hand side is the initial image in the sequence. The images in the right-hand panel are those obtained after 1, 10, 15, 20 and 25 frames have elapsed. Here the sequence is captured at a rate of 10 frames per second. The matches shown are directly from the left-hand frame to the right-hand frame, i.e. no intermediate frames are used. In the case of the dual-step EM algorithm, the matches are all correct in each frame. In the case of the SVD method the method begins to break after 17 frames. In the final example there are no correspondence errors.

7. Conclusions

We have considered three ways of improving the recovery of point correspondences using spectral analysis of the point proximity matrix. The three lines of investigation are
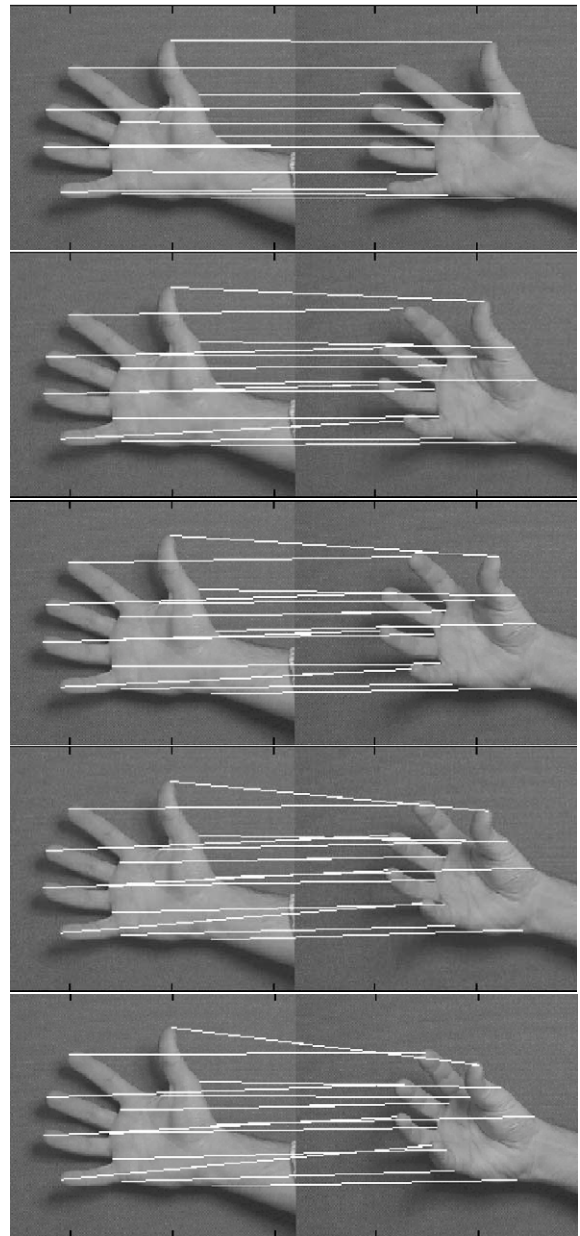
the use of an alternative proximity weighting matrix, the use of robust methods for comparing the eigenvectors of the proximity matrix, and, embedding the correspondence process within the EM algorithm. The first two refinements offer improved performance under point position error, but cannot render the method robust to structural error (i.e. deletion or insertion of points). To overcome structural error, we embed the correspondence process within the EM algorithm.
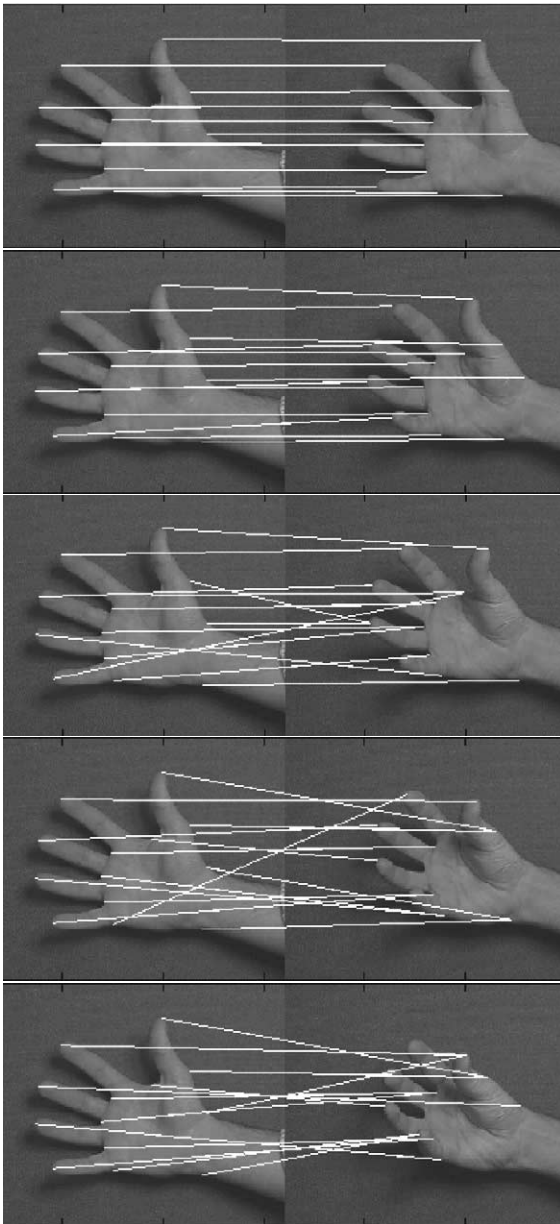
Fig. 11. Matches—pure SVD.

References

[1] F.R.K. Chung, Spectral Graph Theory, CBMS Series, Vol. 92, American Mathematical Society, Providence, RI, 1997.

[2] S. Umeyama, An eigen decomposition approach to weighted graph matching problems, Trans. Pattern Anal. Mach. Intell. 10 (1988) 695–703.

[3] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models: their training and application, CVIU 61 (1) (1995) 38–59.

[4] S. Sclaroff, A.P. Pentland, Modal matching for correspondence and recognition, Trans. Pattern Anal. Mach. Intell. 17 (6) (1995) 545–561.

[5] G.L. Scott, H.C. Longuet-Higgins, An algorithm for associating the features of 2 images, Proc. Roy. Soc. London Ser. B (Biological) 244 (1991) 21–26.

[6] L.S. Shapiro, J.M. Brady, Feature-based correspondence—an eigenvector approach, Image Vision Comput. 10 (1992) 283–288.

[7] M. Pilu, A direct method for stereo correspondence based on singular value decomposition, IEEE Computer Vision and Pattern Recognition Conference, 1997, pp. 261–266.

[8] C. Tomasi, T. Kanade, Shape and motion from image streams under orthography: a factorization method, IJCV 9 (2) (1992) 137–154.

[9] K. Sengupta, K.L. Boyer, Modelbase partitioning using property matrix spectra, Comput. Vision Image Understanding 70 (2) (1998) 177–196.

[10] I.L. Dryden, K.V. Mardia, Statistical Shape Analysis, Wiley, New York, 1998.

[11] H. Sossa, R. Horaud, Model indexing: the graph-hashing approach, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 1992, pp. 811–815.

[12] A. Shokoufandeh, S.J. Dickinson, K. Siddiqi, S.W. Zucker, Indexing using a spectral encoding of topological structure, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1999, pp. 491–497.

[13] S. Sarkar, P. Soundararajan, Supervised learning of large perceptual organization: graph spectral partitioning and learning automata, Trans. PAMI 22 (5) (2000) 504–525.

[14] J. Shi, J. Malik, Normalized cuts and image segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1997.

[15] S. Sarkar, K.L. Boyer, Quantitative measures of change based on feature organization: eigenvalues and eigenvectors, CVIU 71 (1) (1998) 110–136.

[16] W.T. Freeman, P. Perona, A factorisation approach to grouping, European Conference on Computer Vision, 1998, pp. 655–670.

[17] Y. Weiss, Segmentation using eigenvectors: a unifying view, International Conference on Computer Vision, 1999, pp. 975–982.

[18] K. Inoue, K. Urahama, Sequential fuzzy cluster extraction by a graph spectral method, Pattern Recognition Lett. 20 (1999) 699–705.

[19] T. Hofmann, J.M. Buhmann, Pairwise data clustering by deterministic annealing, Trans. Pattern Anal. Mach. Intell. 19 (1) (1997) 1–14.

[20] M. Pelillo, K. Siddiqi, S.W. Zucker, Matching hierarchical structures using association graphs, Trans. Pattern Anal. Mach. Intell. 21 (11) (1999) 1105–1120.

[21] R.C. Wilson, A.D.J. Cross, E.R. Hancock, Structural matching with active triangulations, CVIU 72 (1) (1998) 21–38.

[22] R.C. Wilson, E.R. Hancock, Structural matching by discrete relaxation, IEEE Trans. Pattern Anal. Mach. Intell. 19 (1997) 634–648.

[23] S. Gold, A. Rangarajan, A graduated assignment algorithm for graph matching, Trans. Pattern Anal. Mach. Intell. 18 (4) (1996) 377–388.

[24] S. Gold, A. Rangarajan, C.P. Lu, S. Pappu, E. Mjolsness, New algorithms for 2D and 3D point matching: pose estimation

and correspondence Pattern Recognition 31 (8) (1998) 1019–1031.

[25] A.D.J. Cross, E.R. Hancock, Graph matching with a dual step EM algorithm, IEEE Trans. Pattern Anal. Mach. Intell. 20 (1998) 1236–1253.

[26] M.I. Jordan, R.A. Jacobs, Hierarchical mixtures of experts and the EM algorithm, Neural Comput. 6 (1994) 181–214.

About the Author—MARCO CARCASSONI received his Laurea degree from the University of Padua in Computer Science and Engineering in 1997. Since 1998 he has been working towards a DPhil degree in Computer Vision in the Department of Computer Science at the University of York. He is currently employed as a Research Associate working on an EPSRC project concerned with image retrieval. He has published some 10 papers in image analysis, computer vision and pattern recognition. He is a member of the IEEE.

About the Author—EDWIN HANCOCK studied Physics as an undergraduate at the University of Durham and graduated with honours in 1977. He remained at Durham to complete a Ph.D. in the area of High Energy Physics in 1981. Following this he worked for 10 years as a researcher in the fields of high-energy nuclear physics and pattern recognition at the Rutherford-Appleton Laboratory (now the Central Research Laboratory of the Research Councils). During this period he also held adjunct teaching posts at the University of Surrey and the Open University. In 1991 he moved to the University of York as a lecturer in the Department of Computer Science. He was promoted to Senior Lecturer in 1997 and to Reader in 1998. In 1998 he was appointed to a Chair in Computer Vision.

Prof. Hancock now leads a group of some 15 faculty, research staff and Ph.D. students working in the areas of computer vision and pattern recognition. His main research interests are in the use of optimisation and probabilistic methods for high and intermediate level vision. He is also interested in the methodology of structural and statistical pattern recognition. He is currently working on graph-matching, shape-from-X, image data-bases and statistical learning theory. His work has found applications in areas such as radar terrain analysis, seismic section analysis, remote sensing and medical imaging. Professor Hancock has published some 60 journal papers and 200 refereed conference publications. He was awarded the Pattern Recognition Society medal in 1991 for the best paper to be published in the journal Pattern Recognition. The journal also awarded him an outstanding paper award in 1997.

Prof. Hancock has been a member of the Editorial Boards of the journals IEEE Transactions on Pattern Analysis and Machine Intelligence, and, Pattern Recognition. He has also been a guest editor for special editions of the journals Image and Vision Computing and Pattern Recognition, and he is currently a guest editor of a special edition of IEEE Transactions on Pattern Analysis and Machine Intelligence devoted to energy minimisation methods in computer vision. He has been on the programme committees for numerous national and international meetings. In 1997 he established a new series of international meetings on energy minimisation methods in computer vision and pattern recognition. He was awarded a Fellowship of the International Association for Pattern Recognition in 2000.