Alignment and Correspondence Using Singular Value Decomposition

B. Luo[1,2] and E.R. Hancock[1]

[1] Department of Computer Science,
University of York,York YO10 5DD, UK.
[2] Anhui University, P.R. China
{luo,erh}@minster.cs.york.ac.uk

Abstract. This paper casts the problem of point-set alignment and correspondence into a unified framework. The utility measure underpinning the work is the cross-entropy between probability distributions for alignment and assignment errors. We show how Procrustes alignment parameters and correspondence probabilities can be located using dual singular value decompositions. Experimental results using both synthetic and real images are given.

1 Introduction

Point pattern matching is a problem of pivotal importance in computer vision that continues to attract considerable interest. The problem may be abstracted as either alignment or correspondence. Alignment involves explicitly transforming the point positions under a predefined geometry so as to maximise a measure of correlation. Examples here include Procrustes normalisation [8], affine template matching [20] and deformable point models [2]. Correspondence, on the other hand, involves recovering a consistent arrangement of point assignment labels. The correspondence problem can be solved using a variety of point assignment [13,15] and graph-matching [21,6,1,9] algorithms.

The problem of point pattern matching has attracted sustained interest in both the vision and statistics communities for several decades. For instance, Kendall [8] has generalised the process to projective manifolds using the concept of Procrustes distance. Ullman [16] was one of the first to recognise the importance of exploiting rigidity constraints in the correspondence matching of point-sets. Recently, several authors have drawn inspiration from Ullman's ideas in developing general purpose correspondence matching algorithms using the Gaussian weighted proximity matrix. For instance Scott and Longuet-Higgins [13] locate correspondences by finding a singular value decomposition of the inter-image proximity matrix. Shapiro and Brady [14], on the other hand, match by comparing the modal eigenstructure of the intra-image proximity matrix. In fact these two ideas provide some of the basic groundwork on which the deformable shape models of Cootes *et al* [2] and Sclaroff and Pentland [12] build. This work on the co-ordinate proximity matrix is closely akin to that of Umeyama [17] who shows

F.J. Ferri et al. (Eds.): SSPR&SPR 2000, LNCS 1876, pp. 226–235, 2000.

how point-sets abstracted in a structural manner using weighted adjacency graphs can be matched using an eigen-decomposition method. These ideas have been extended to accommodate parametererised transformations [18] which can be applied to the matching of articulated objects [19]. More recently, there have been several attempts at modelling the structural deformation of point-sets. For instance, Amit and Kong [1] have used a graph-based representation (graphical templates) to model deforming two-dimensional shapes in medical images. Lades *et al* [9] have used a dynamic mesh to model intensity-based appearance in images.

In a recent paper we developed a unified statistical framework for alignment and correspondence [3]. The motivation for the work was that the dichotomy normally drawn between the two processes overlooks considerable scope for synergistic interchange of information. In other words, there must always be bounds on alignment before correspondence analysis can be attempted, and vice versa. Our approach in developing the new point-pattern matching method was to embed constraints on the spatial arrangement of correspondences within an EM algorithm for alignment parameter recovery. This process has many features reminiscent of Jordan and Jacob's hierarchical mixture of experts algorithm [7]. The observation underpinning this paper is that although the method proved effective it fails to put the alignment and correspondence processes on a symmetric footing. The relational constraints were simply used to gate the contributions to the log-likelihood function for the alignment errors.

The idea underpinning this paper is to provide a new framework for the maximum likelihood matching of point-sets which allows a symmetric linkage between alignment and correspondence. Specifically, we aim to realise interleaved iterative steps which communicate via an integrated utility measure. The utility measure is the cross-entropy between the probability distributions for alignment and correspondence. By casting the cross-entropy in terms of matrices, we realise optimisation via dual singular value decompositions. The first of these transforms the point set positions so as to locate an alignment that maximises the weighted correlation between the point-set co-ordinates. The second singular value decomposition updates the set of correspondence probabilities that maximise the weighted correlation between the edge-sets of the adjacency graphs for the point-sets. These processes are interleaved and iterated to convergence.

2 Point-Sets

Our goal is to recover the Procrustes normalisation that best aligns a set of image feature points $\mathbf{w}$ with their counterparts in a model $\mathbf{z}$. In order to do this, we represent each point in the image data set by a co-ordinate position vector $\boldsymbol{w}_i = (x_i, y_i)^T$ where i is the point index. In the interests of brevity we will denote the entire set of image points by $\mathbf{w} = \{\boldsymbol{w}_i, \forall i \in \mathcal{D}\}$ where $\mathcal{D}$ is the point index-set. The corresponding fiducial points constituting the model are similarly represented by $\mathbf{z} = \{\boldsymbol{z}_j, \forall j \in \mathcal{M}\}$ where $\mathcal{M}$ denotes the index-set for the model feature-points $\boldsymbol{z}_j$.

Later on we will show how the two point-sets can be aligned using singular value decomposition. In order to establish the required matrix representation of the alignment process, we construct two co-ordinate matrices from the point position vectors. The data-points are represented by the matrix $D = (\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad .... \quad \boldsymbol{w}_{|D|})$ whose columns are the co-ordinate position vectors. The corresponding point-position matrix for the model is $M = (\boldsymbol{z}_1 \quad \boldsymbol{z}_2 \quad .... \quad \boldsymbol{z}_{|M|})$.

One of our goals in this paper is to exploit structural constraints to improve the recovery of alignment parameters from sets of feature points. To this end we represent point adjacency using a neighbourhood graph. There are many alternatives including the N-nearest neighbour graph, the Delaunay graph, the Gabriel graph and the relative neighbourhood graph. Because of its well documented robustness to noise and change of viewpoint, we adopt the Delaunay triangulation as our basic representation of image structure [4]. We establish Delaunay triangulations on the data and the model, by seeding Voronoi tessellations from the feature-points.

The process of Delaunay triangulation generates relational graphs from the two sets of point-features. More formally, the point-sets are the nodes of a data graph $G_D = \{\mathcal{D}, E_D\}$ and a model graph $G_M = \{\mathcal{M}, E_M\}$. Here $E_D \subseteq \mathcal{D} \times \mathcal{D}$ and $E_M \subseteq \mathcal{M} \times \mathcal{M}$ are the edge-sets of the data and model graphs. Later on we will cast our optimisation process into a matrix representation. Here we use the notation $\hat{E}_D(i, i')$ to represent the elements of the adjacency matrix for the data graph; the elements are unity if $i = i'$ or if (i, i') is an edge and are zero otherwise. We represent the state of correspondence between the two graph using the function $f : \mathcal{D} \rightarrow \mathcal{M}$ from the nodes of the model graph onto the nodes of the data-graph.

3 Dual Step Matching Algorithm

We characterise the matching problem in terms of separate probability distributions for alignment and correspondence. In the case of alignment, the distribution models the registration errors between the data point positions and their counterparts in the model under Procrustes alignment. The correspondence process on the other hand captures the consistency of the pattern of matching assignments to the graph representing the point-sets. The set of assignments is represented by the function $f : \mathcal{D} \rightarrow \mathcal{M}$. Suppose that $P_{i,j}^{(n)}$ is the probability that node i from the data graph is in alignment with node j from the model graph at iteration n. Similarly, $Q_{i,j}^{(n)}$ is the probability that node i is in correspondence with node j. Further suppose that $p_{i,j}^{(n)} = p(\underline{\mathrm{w}}_i^{(n)}|\underline{z}_j)$ is the probability distribution for the alignment error between the nodes i and j under the Procrustes alignment at iteration n. The distribution of the correspondence errors associated with the assignment function $f^{(n)}$ at iteration n is $q_{i,j}^{(n)}$. With these ingredients the utility measure which we aim to maximise in the dual alignment and correspondence step is

$$\mathcal{E} = \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} \left[Q_{i,j}^{(n)} \ln p_{i,j}^{(n+1)} + P_{i,j}^{(n)} \ln q_{i,j}^{(n+1)} \right] \tag{1}$$

In other words, the two processes interact via a symmetric expected log-likelihood function. The correspondence probabilities weight contributions to the expected log-likelihood function for the alignment errors, and vice-versa. In our previous work, we showed how the first term arises through the gating of the log-likelihood function of the EM algorithm [3].

The alignment point positions and correspondence matches are recovered via the dual maximisation equations

$$D^{(n+1)} = \arg\max_{D} \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} Q_{i,j}^{(n)} \ln p_{i,j}^{(n+1)} \tag{2}$$

and

$$f^{(n+1)} = \arg\max_{f} \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} P_{i,j}^{(n)} \ln q_{i,j}^{(n+1)} \tag{3}$$

3.1 Alignment

To develop a useful alignment algorithm we require a model for the measurement process. Here we assume that the observed position vectors, i.e. $\boldsymbol{w}_i$ are derived from the model points through a Gaussian error process. Suppose that the revised estimates of the data-point position matrix $D^{(n)}$ is $\boldsymbol{w}_i^{(n)}$. According to our Gaussian model of the alignment errors,

$$p(\boldsymbol{w}_i^{(n)}|\boldsymbol{z}_j) = \frac{1}{2\pi\sqrt{|\Sigma|}} \cdot \exp\left[-\frac{1}{2} (\boldsymbol{z}_j - \boldsymbol{w}_i^{(n)})^T \Sigma^{-1} (\boldsymbol{z}_j - \boldsymbol{w}_i^{(n)}) \right] \tag{4}$$

where Σ is the variance-covariance matrix for the point measurement errors. Here we assume that the position errors are isotropic, in other words the errors in the x and y directions are identical and uncorrelated. As a result we write $\Sigma = \sigma^2 I_2$ where I_2 is the 2x2 identity matrix. With this model, maximisation of the expected log-likelihood function $\mathcal{E}_a$ reduces to minimising the weighted square error measure

$$\mathcal{F}_a = \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} Q_{i,j}^{(n)} (\boldsymbol{z}_j - \boldsymbol{w}_i^{(n+1)})^T (\boldsymbol{z}_j - \boldsymbol{w}_i^{(n+1)}) \tag{5}$$

We would like to recover the maximum likelihood alignment parameters by applying an rigid transformation to the two point-sets. We recover the required parameter matrix by performing singular value decomposition of a point-correspondence matrix. In order to develop the necessary formalism, we rewrite the weighted squared error criterion using a matrix representation. Suppose that

$Q^{(n)}$ is the data-responsibility matrix whose elements are the *a posteriori* correspondence probabilities $Q_{i,j}^{(n)}$. With this notation the quantity $\mathcal{F}_a$ can be expressed in the following matrix form

$$\mathcal{F}_a = Tr[M^T Q^{(n)} M] - 2Tr[D^{(n+1)} Q^{(n)} M^T] + Tr[D^{(n+1)T} Q^{(n)} D^{(n+1)}] \quad (6)$$

Since the first and third terms of this expression do not depend on the alignment of the point-sets, the transformation matrix satisfies the condition

$$D^{(n+1)} = \arg\max_{\hat{D}} Tr[\hat{D} Q^{(n)} M^T] \quad (7)$$

The Procrustes alignment of the points can be thought of as maximising a weighted measure of overlap or correlation between the point-sets.

The required maximisation can be performed using a singular value decomposition. The procedure is as follows. The matrix $DQ^{(n)}M^T$ is factorised into a product of three new matrices U, V and Δ, where Δ is a diagonal matrix whose elements are either zero or positive, and U and V are orthogonal matrices. The factorisation is as follows $D^{(n+1)}Q^{(n)}M^T = U\Delta V^T$.

The matrices U and V define a rotation matrix Θ which aligns the principal component directions of the point-sets M and D. The rotation matrix is equal to $\Theta = VU^T$.

With the rotation matrix to hand we can find the Procrustes alignment which maximises the correlation of the two point sets. The procedure is to first bring the centroids of the two point-sets into correspondence. Next the data points are scaled to that they have the same variance as those of the model. Finally, the scaled and translated point-sets are rotated so that their correlation is maximised.

To be more formal the centroids of the two point-sets are $\mu_D^{(n)} = E(w_i^{(n)})$ and $\mu_M = E(z_j)$. The corresponding covariance matrices are $\Sigma_D^{(n)} = E((w_i^{(n)} - \mu_D^{(n)})(w_i^{(n)} - \mu_D^{(n)})^T)$ and $\Sigma_M = E((z_j - \mu_M^{(n)})(z_j - \mu_M^{(n)})^T)$.

With ingredients the update equation for re-aligning the data-points is

$$\boldsymbol{w}_i^{(n+1)} = \mu_M + \frac{Tr\Sigma_M}{Tr\Sigma_D} VU^T(\boldsymbol{w}_i^{(n)} - \mu_D^{(n)}) \quad (8)$$

Finally, we update the *a posteriori* matching probabilities by substituting the revised position vector 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

$$P_{i,j}^{(n)} = \frac{p(\boldsymbol{w}_i^{(n)}|\boldsymbol{z}_j)}{\sum_{j' \in \mathcal{M}} p(\boldsymbol{w}_i^{(n)}|\boldsymbol{z}_{j'})} \quad (9)$$

It is worth pausing to consider the relationship between the point correlation measure developed in this paper and those exploited elsewhere in the literature on point pattern matching. The quantity DM^T is simply the standard measure

of overlap that is minimised in the work on least-squares alignment [18]. The matrix Q, on the other hand, plays the role of the correspondence matrix used by Scott and Longuet-Higgins [13]. So, the utility measure delivered by the cross entropy plays a synergistic role with the correspondence point proximity matrix weights the least-squares criterion.

3.2 Correspondences

The correspondences are recovered via maximisation of the quantity

$$\mathcal{E}_c = \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} P_{i,j}^{(n)} \ln q_{i,j}^{(n+1)} \tag{10}$$

Suppose that $V_D(i) = \{i' | (i, i') \in E_D\}$ represents the set of nodes connected to the node i by an edge in the graph with edge-set E_D. Furthermore, let us introduce a set of assignment variables that convey the following meaning

$$s_{i,j}^{(n)} = \begin{cases} 1 & \text{if } f^{(n)}(i) = j \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

In a recent study [5], we have shown that the probability distribution for the assignment variables is

$$q_{i,j}^{(n)} = K \exp\left[-k_e \sum_{i' \in V_D(i)} \sum_{j' \in V_M(j)} (1 - s_{i',j'}^{(n+1)}) \right] \tag{12}$$

where K and k_e are constants. With this distribution to hand, the correspondence assignment step reduces to one of maximising the quantity

$$\mathcal{F}_c = \sum_{i \in \mathcal{D}} \sum_{j \in \mathcal{M}} \sum_{i' \in \mathcal{D}} \sum_{j' \in \mathcal{M}} \hat{E}_D(i, i') \hat{E}_M(j, j') P_{i,j}^{(n)} s_{i',j'}^{(n+1)} \tag{13}$$

where $\hat{E}_D(i, i')$ and $\hat{E}_M(j, j')$ are the elements of the adjacency matrices for the data and model graphs. In more compact notation, the updated matrix of correspondence indicators $S^{(n+1)}$ satisfies the condition

$$S^{(n+1)} = \arg\max_S Tr[\hat{E}_D^T P^{(n)} \hat{E}_M S^T] \tag{14}$$

where $P^{(n)}$ is a matrix whose elements are the alignment probability $P_{i,j}^{(n)}$. In other words, the utility measure gauges the degree of correlation between the edge-sets of the two graphs under the permutation structure induced by the alignment and correspondence probabilities. Following Scott and Longuet-Higgins [13] we recover the matrix of assignment variables that maximises $\mathcal{F}_c$ by performing the singular value decomposition $\hat{E}_D^T P^{(n)} \hat{E}_M = V \Delta U^T$, where Δ is again a diagonal matrix and U and V are orthogonal matrices. The matrices U and V are used to compute an assignment matrix $R^{(n+1)} = V U^T$. To compute

the associated matrix of correspondence probabilities, $Q^{(n+1)}$, we perform row normalisation on $R^{(n+1)}$. As a result

$$Q_{i,j}^{(n)} = \frac{R_{i,j}^{(n+1)}.}{\sum_{j \in \mathcal{M}} R_{i,j}^{(n+1)}} \tag{15}$$

This is clearly simplistic and violates symmetry. In our future work we plan to improve the algorithm to include Sinckhorn normalisation and slack variables for unmatchable nodes along the lines of Gold and Rangarajan [6].

4 Experiments

In this section, we provide some experimental evaluation of the new unified approach to correspondence and alignment. Here, we use both synthetic point-sets and real images.

4.1 Sensitivity Study

To evaluate the robustness of the new approach, we furnish a sensitivity study. This compares the new iterative alignment method with the following three alternatives:

- The first method(Refered to as "Weight+SVD") is similar to that of Scott and Longuet-Higgins. This performs the singular value decomposition $W^{(0)} = U_S \Delta_S V_S^T$ on the initial inter-image weight matrix. Suppose that $\hat{\Delta}_S$ is the matrix obtained by setting the diagonal elements of Δ_S to unity, then the Scott and Longuet-Higgins algorithm delivers an updated matrix of correspondence weights $\hat{W} = U_S \hat{\Delta}_S V_S^T$. The updated weight matrix can be used to align the point-sets using the method outlined earlier in this paper.
- The second algorithm(Refered to as "Single SVD") performs the singular value decomposition $DM^T = U \Delta V^T$ to find the rotation matrix $\Theta = VU^T$ that maximises the unweighted point correlation $Tr[DM^T]$.
- The third method(Refered to as "PCA") is based upon aligning and scaling in the principal component axes of the two point-sets.

Figure 1a shows the RMS error as a function of the standard deviation of the point position error. The main point to note from this plot is that for all four algorithms the RMS error increases linearly with the noise standard deviation. However, for the new algorithm (EM+Weight+SVD-shown as circle points), the rate of increase of the RMS error is much lower than the remaining three algorithms. In other words, the new algorithm gives more accurate alignments.

Figure 1b shows the fraction of points in correct correspondence as a function of the fraction of added clutter. The main point to note for this plot is that the new method (EM+Weight+SVD - shown as circles) is considerably more accurate in locating correspondences. Moreover, the two SVD-based methods perform only marginally better than the PCA alignment,

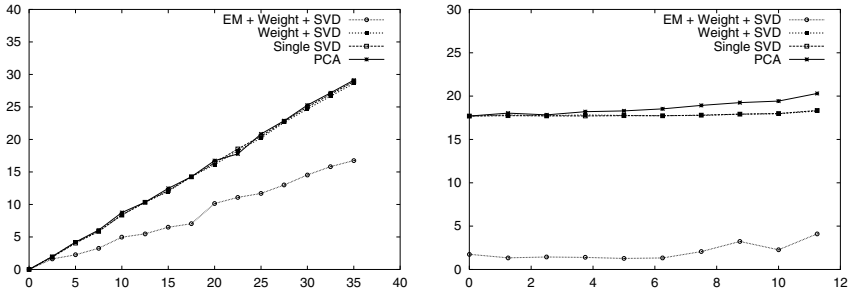

Fig. 1. Sensitivity study. Left:Alignment error as a function of noise-variance on the point-sets. Right:Alignment error as a function of the fraction of structural error

We have evaluated the noise sensitivity of the algorithm on synthetic point-sets and compared it with the quadratic assignment method of Gold and Rangarajan [11] and the Scott and Longuet-Higgins algorithm. The point sets have been subjected to affine deformation, random measurement error(positional jitter) and contamination from added clutter.

Figure 2 shows the fraction of correct correspondences as a function of the fraction of added clutter. The method outperforms that of the quadratic assignment algorithm and the Scott and Longuet-Higgins algorithm. The onset of significant correspondence error occurs when the fraction of clutter exceeds 30%.

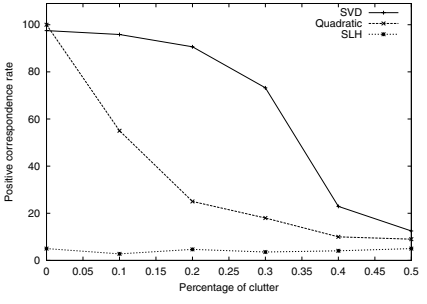

Fig. 2. Positive correspondence rate as a function of percentage of clutters for rigid transformed point set with clutters.

4.2 Real-World Data

We have evaluated the algorithm on matching corners detected in real-world images. The corner detector used in our studies is described in the recent paper by by Luo, Cross and Hancock[10]. We use Delaunay graphs to represent the structural arrangement of the corners. Figure 4 shows the correspondences between the corners as lines between the two images. After checking by hand, the fraction of correct correspondences is 77%. Figure 5 shows the iterations of the alignment process. The process converges after 10 iterations and the alignment is qualitatively good.

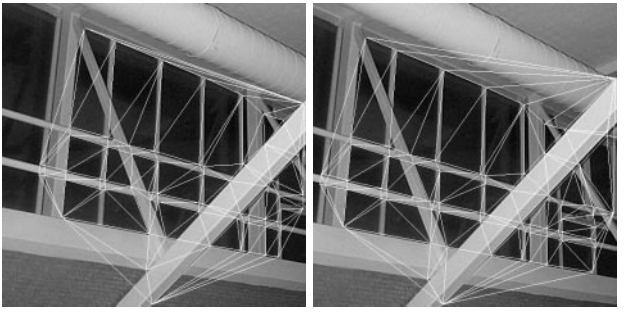

Fig. 3. Test images overlayed with Delaunay graph

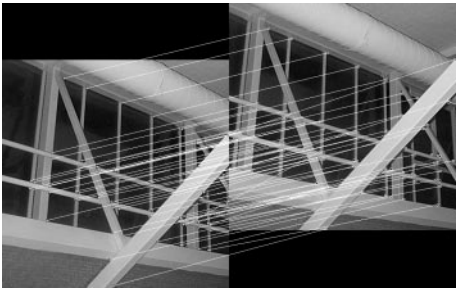

Fig. 4. Correspondences

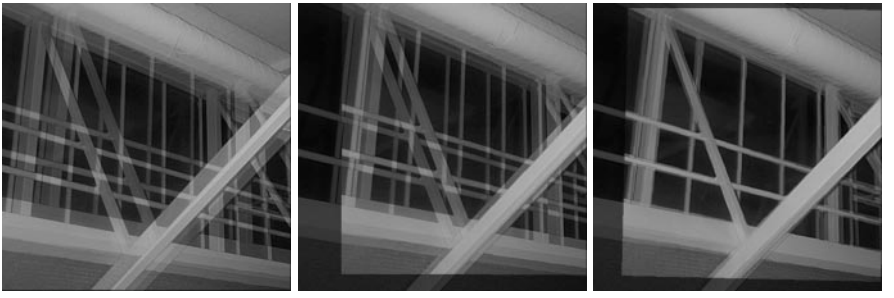

Fig. 5. Alignment results. Left:Overlayed original images. Middle: 1 Iteration. Right: 10 Iterations

5 Conclusions

In conclusion, we have shown how the processes of point-set alignment and correspondence analysis can be unified using a symmetric entropy. By drawing on a Gaussian model of point position errors and an exponential model of correspondence assignment errors, we are able to cast the two problems as maximisation of weighted correlation measures. In both cases the point matches can be recovered using singular value decomposition. Our new measures of point-set similarity naturally combine the ideas already developed by Scott and Longuet-Higgins and Umeyama in a single statistical utility measure. An experimental study reveals

that the proposed method is outperforms that of Scott and Longuet-Higgins in terms of its ability to recover from contaminating clutter and positional error in the point-sets.

References

1. Y. Amit and A. Kong. Graphical Templates for Model Registration. *IEEE PAMI*, 18:225–236, 1996.
2. T.F Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active Shape Models - Their Training and Application. *CVIU*, 61:38–59, 1995.
3. A.D.J. Cross and E.R. Hancock. Graph Matching with a Dual-Step EM Algorithm. *IEEE PAMI*, 20(11):1236–1253, 1998.
4. O.D. Faugeras, E. Le Bras-Mehlman, and J-D. Boissonnat. Representing Stereo Data with the Delaunay Triangulation. *Artificial Intelligence*, 44:41–87, 1990.
5. A.M. Finch, R.C. Wilson, and E.R. Hancock. Matching delaunay graphs. *Pattern Recognition*, 30(1):123–140, January 1997.
6. S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *IEEE PAMI*, 18(4):377–388, 1996.
7. M.I. Jordan and R.A. Jacobs. Hierarchical mixture of experts and the EM algorithm. *Neural Computation*, 6:181–214, 1994.
8. D.G. Kendall. Shape Manifolds: Procrustean metrics and complex projective spaces. *Bulletin of the London Mathematical Society*, 16:81–121, 1984.
9. M. Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. von der Maalsburg, R.P. Wurtz, and W.Konen. Distortion-invariant object-recognition in a dynamic link architecture. *IEEE Transactions on Computers*, 42:300–311, 1993.
10. B. Luo, A.D.J. Cross, and E.R. Hancock. Corner detection via topographic analysis of vector potential. *Pattern Recognition Letters*, 20:635–650, 1999.
11. A. Rangarajan and E. D. Mjolsness. A lagrangian relaxation network for graph matching. *IEEE Transactions on Neural Networks*, 7(6):1365–1381, November 1996.
12. S. Sclaroff and A.P. Pentland. Modal Matching for Correspondence and Recognition. *IEEE PAMI*, 17:545–661, 1995.
13. 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(1309):21–26, 1991.
14. L.S. Shapiro and J. M. Brady. Feature-based Correspondence - An Eigenvector Approach. *Image and Vision Computing*, 10:283–288, 1992.
15. L.S. Shapiro and J. M. Brady. Rejecting Outliers and Estimating Errors in an Orthogonal-regression Framework. *Phil. Trans. Roy. Soc. A*, 350:403–439, 1995.
16. S. Ullman. The Interpretation of Visual Motion. *MIT Press*, 1979.
17. S. Umeyama. An Eigen Decomposition Approach to Weighted Graph Matching Problems. *IEEE PAMI*, 10:695–703, 1988.
18. S. Umeyama. Least Squares Estimation of Transformation Parameters between Two Point sets. *IEEE PAMI*, 13(4):376–380, 1991.
19. S. Umeyama. Parameterised Point Pattern Matching and its Application to Recognition of Object Families. *IEEE PAMI*, 15:136–144, 1993.
20. M. Werman and D. Weinshall. Similarity and Affine Invariant Distances between 2D Point Sets. *IEEE Trans. on PAMI*, 17(8):810–814, 1995.
21. R.C. Wilson and E.R. Hancock. Structural matching by discrete relaxation. *IEEE PAMI*, 19(6):634–648, June 1997.