PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

An Algorithm for Associating the Features of Two Images

Guy L. Scott and H. Christopher Longuet-Higgins

Proc. R. Soc. Lond. B 1991 **244**, 21-26
doi: 10.1098/rspb.1991.0045

An algorithm for associating the features of two images

GUY L. SCOTT AND H. CHRISTOPHER LONGUET-HIGGINS†

Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, U.K.

SUMMARY

In this paper we describe an algorithm that operates on the distances between features in the two related images and delivers a set of correspondences between them. The algorithm maximizes the inner product of two matrices, one of which is the desired 'pairing matrix' and the other a 'proximity matrix' with elements $\exp(-r_{ij}^2/2\sigma^2)$, where r_{ij} is the distance between two features, one in each image, and σ is an adjustable scale parameter. The output of the algorithm may be compared with the movements that people perceive when viewing two images in quick succession, and it is found that an increase in σ affects the computed correspondences in much the same way as an increase in interstimulus interval alters the perceived displacements. Provided that σ is not too small the algorithm will recover the feature mappings that result from image translation, expansion or shear deformation – transformations of common occurrence in image sequences – even when the displacements of individual features depart slightly from the general trend.

1. INTRODUCTION

A central problem in the theory of vision (Wertheimer 1912; Marr 1976; Ullman 1979) is that of establishing a correspondence between the features of two related images such as the members of a stereo pair or successive frames in a motion sequence. If an individual feature is sufficiently distinctive, there may be no problem in tracking it through a sequence of images, but when the features are small and unstructured a conflict arises between two principles, both of them soundly based in visual experience. The first principle requires that – other things being equal (not an easy condition to define) – a match across a shorter distance is to be favoured. This 'principle of proximity' is sufficiently influential in human vision to disrupt the perception of global rigidity in, for example, Ullman's 'broken wheel' demonstration (Ullman 1979). However, it is no solution to the correspondence problem simply to associate each point in one image with its nearest neighbour in the other. A little doodling will reveal that this usually gives rise to many-to-one correspondences between features, and to results that differ according to which is chosen as the 'reference' image.

To account for the movements that people perceive when related patterns are viewed in quick succession, one must suppose the principle of proximity to operate within the limits imposed by a 'principle of exclusion', that militates against many-to-one feature correspondences. In figure 1 the lines linking the circles and the crosses indicate the 'coherent' displacements generally perceived when the two patterns are shown alternately as a 'movie'. This one-to-one mapping is quite different from the 4:1 and 1:4 'nearest neighbour' mappings implied by the principle of proximity taken on its own.

Here we propose a simple algorithm that incorporates both the principle of proximity and the principle of exclusion. Our algorithm resembles Ullman's minimal mapping theory in maximizing the inner product of a given matrix G and a pairing matrix P; but whereas Ullman introduces the exclusion principle as an explicit constraint in a linear programming exercise, we find it to emerge naturally from the requirement that the rows of P (possibly fewer than the columns) be mutually orthogonal. In brief, our pairing matrix P is that orthogonal matrix that maximizes the inner product $P{:}G$, where G is a matrix of 'proximities' between the features in one image and those in the other. Once P has been computed, the feature correspondences follow immediately.

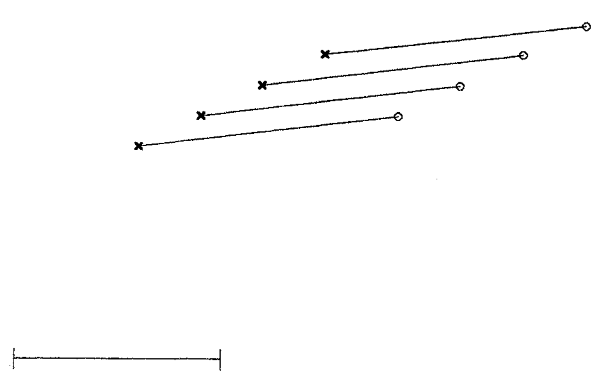

Figure 1. An example of the sort of 'coherent' movement that is generally perceived when the inter-stimulus interval is not too short.

† Current address: Centre for Research on Perception and Cognition, Laboratory of Experimental Psychology, University of Sussex, Falmer, Brighton BN1 9QG, U.K.

Proc. R. Soc. Lond. B (1991) **244**, 21–26
Printed in Great Britain

2. THE ALGORITHM

Let I and J be two related images, containing m and n features respectively. We regard the images as lying in the same plane, so that there is a well-defined distance r_{ij} between feature I_i $(i = 1, 2, \ldots, m)$ and feature J_j $(j = 1, 2, \ldots, n)$. In our diagrams the features I_i are represented by crosses and the J_j by circles. We assume, without loss of generality, that $m \leqslant n$.

The first stage in our algorithm is to represent the 'proximities' between features in I and features in J by a rectangular $m \times n$ matrix G with elements

$$G_{ij} = \exp(-r_{ij}^2/2\sigma^2),$$

where σ is an appropriate unit of distance. The Gaussian form of G_{ij} has various useful properties. First, it is analytic not only in the distance r_{ij} but also in the coordinates of the points I_i and J_j. Secondly, it is provably positive definite (see Appendix 1) when the two images are identical, a fact that has useful implications for the study of its analytic behaviour. Thirdly, the distance σ provides a convenient variable with which to study the difference between 'large' and 'small' displacements; and finally, G_{ij} decreases monotonically with distance, from 1 for neighbouring points to 0 for points that are a long way apart. (In Ullman's theory the nearest analogue to G_{ij} is his 'cost' function $q(v)$, a monotonically increasing function of distance.)

The next stage in our algorithm is to submit the proximity matrix G to singular-value decomposition (see, for example, Strang 1988); that is, we express it as a product of the form

$$G = TDU,$$

where T and U are orthogonal matrices of dimension m and n respectively, and D is a non-negative diagonal matrix. (Opinions differ as to the neatest way of representing the singular-value decomposition of a rectangular matrix. Here we take D to be a matrix of the same shape as G, satisfying the constraints $D_{ij} = 0$ $(i \neq j)$, $D_{ii} > 0$. It is possible for one or more of the singular values D_{ii} to vanish, but we ignore such cases for the time being. Finally, if m is less than n, only the first m rows of the matrix U have any significance; the others play no part in the singular value decomposition of G, or in our computations.)

Finally, we convert the matrix D into another $m \times n$ matrix E by replacing every diagonal element D_{ii} by the number 1, and obtain another 'orthogonal' matrix

$$P = TEU$$

of the same shape as the original proximity matrix G. (Here and later the term 'orthogonal' is extended to rectangular $m \times n$ matrices whose rows are mutually orthogonal even though their columns are not.)

The rows of P, like those of G, index the features in the first image, and its columns those in the second. The element P_{ij} indicates the extent of pairing between features I_i and J_j. If P_{ij} is the greatest element in row i but not the greatest in column j, then we may regard I_i as competing unsuccessfully for partnership with J_j; similar remarks apply if P_{ij} is the greatest element in its column but not in its row. But if P_{ij} is *both* the greatest element in its row *and* the greatest element in its column then we regard those two features as being in 1:1 correspondence with one another; such correspondences are those implied by the full lines in the figures we shall soon encounter.

3. AN EXTREMUM PRINCIPLE

We now show that the orthogonal matrix P obtained in the manner described is the one that 'correlates best' with G, in the sense of maximizing the inner product

$$P{:}G = \Sigma_i \Sigma_j P_{ij} G_{ij} = \mathrm{trace}\,(P^T G),$$

where P^T denotes the transpose of P.

Let F denote any orthogonal $m \times n$ matrix. Then, because all the diagonal elements of D are non-negative, trace $(F^T D)$ attains its maximum value when F is the unit matrix E defined earlier. To proceed we note that

$$D = T^T G U^T$$

and use the fact that the trace of a product of matrices is invariant under a cyclic permutation of its factors. Writing

$$\mathrm{trace}\,(F^T D) = \mathrm{trace}\,(F^T T^T G U^T) = \mathrm{trace}\,(U^T F^T T^T G) = \mathrm{trace}\,(Q^T G),$$

where Q is defined as

$$Q = TFU,$$

we infer that the orthogonal $(m \times n)$ matrix Q that maximizes trace $(Q^T G)$ is none other than the matrix

$$P = TEU.$$

It is illuminating to see what happens when the algorithm is applied to the special case in which the second image is identical with the first. The proximity matrix G will then be square, symmetric and positive definite: square because the two sets of features are equal in number; symmetric because the distance between features I_i and J_j equals the distance between features I_j and J_i; and positive definite because of the Gaussian form of its elements (see Appendix 1). As a result, the orthogonal matrices T and U in the singular-value decomposition of G will be mutually inverse, each being the transpose of the other, and the pairing matrix P reduces to the unit matrix E

$$P = TEU = TU = E,$$

in which every non-zero element is both the greatest in its row and the greatest in its column.

As the two images begin to depart from one another, the proximity matrix G ceases to be fully symmetric, and the off-diagonal elements of P acquire non-zero values; but the divergence between the images has to be quite substantial before the originally diagonal character of P, and the implied 1:1 correspondence between their features, is totally obliterated.

The pairing matrix P incorporates the principle of exclusion by virtue of its orthogonality. (Theoretical chemists may detect echoes of the Pauli exclusion principle, and its application to electron orbitals.) The fact that the squares of the elements in each row of P must add up to 1 implies that a given feature I_i cannot be strongly associated with more than one feature J_j, though it may be weakly associated with several; the mutual orthogonality of the rows tends to keep different features in the first image from becoming closely associated with the same feature in the second image.

Before reporting some of our computational experiments with the algorithm we should discuss the probability that one or more of the elements D_{ii}, the singular values of G, is equal to zero. This happens when, for example, the features I_1 and I_2 lie at opposite corners of a square, and the features J_1 and J_2 occupy the other two corners. Then all the elements of G are equal, by symmetry, to some number α, and its singular-value decomposition yields a D matrix with diagonal elements 2α and 0, and T and U matrices that are mutual inverses. But now there is no constraint involving the second row of U or the second column of T; the elements of either can be reversed in sign without affecting the magnitude of $P{:}G = \text{trace}\,(T^T G U^T)$. The result is an ambiguity in the form of P; it might be either

$$\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \quad \text{or} \quad \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}\text{'}$$

matrices that represent the mappings associated with the two pairs of opposite sides of the square. A slight shortening of either pair of edges results, not unnaturally, in one of these mappings being preferred to the other.

Although in the above case $P{:}G$ is maximized by either of two permutation matrices, this is not necessarily true of image pairs that hover between two alternative mappings. If, for example, the circles and the crosses comprise the two triangles of a regular hexagram, P is uniquely determined by the maximization of $P{:}G$; the alternative pairings arise from the high symmetry of P itself

$$\begin{matrix} -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{matrix}$$

which permits two distinct optimal mappings: rotation of either triangle through $\pm\pi/3$.

4. COMPUTATIONAL RESULTS

Figures 2–8 illustrate the performance of the algorithm on various pairs of images. The crosses mark the positions of features in the first image, the circles those in the second. A full line joining a cross and a circle indicates a pair of features for which P_{ij} is the greatest element in both its row and its column; a dotted line indicates that one or other of the two features is more strongly associated with a third. The value of σ is indicated by the line segment at the foot of each diagram. In most of the diagrams the two images comprised equal numbers of features, but in some of the runs a few extra features, indicated by filled squares, were added to the first image, to see how much disruption they created.

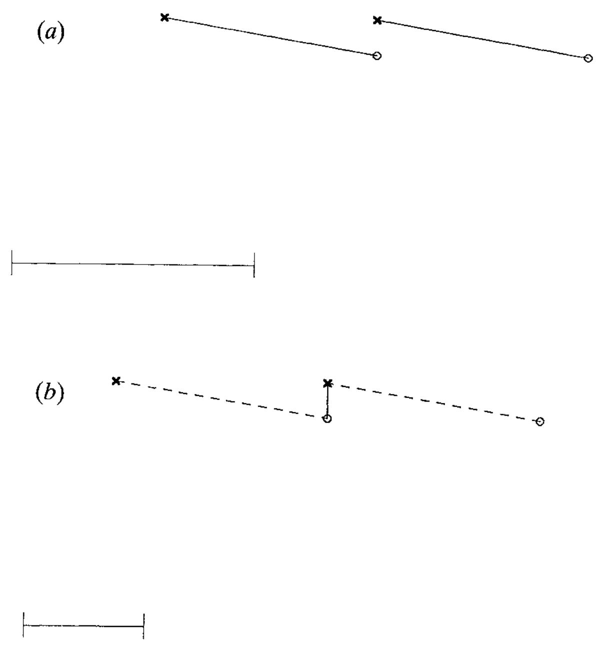

Figure 2. (*a*) Result of applying our algorithm to Ternus's configuration with 'long-range' σ. (*b*) Result of applying our algorithm to Ternus's configuration with 'short-range' σ.

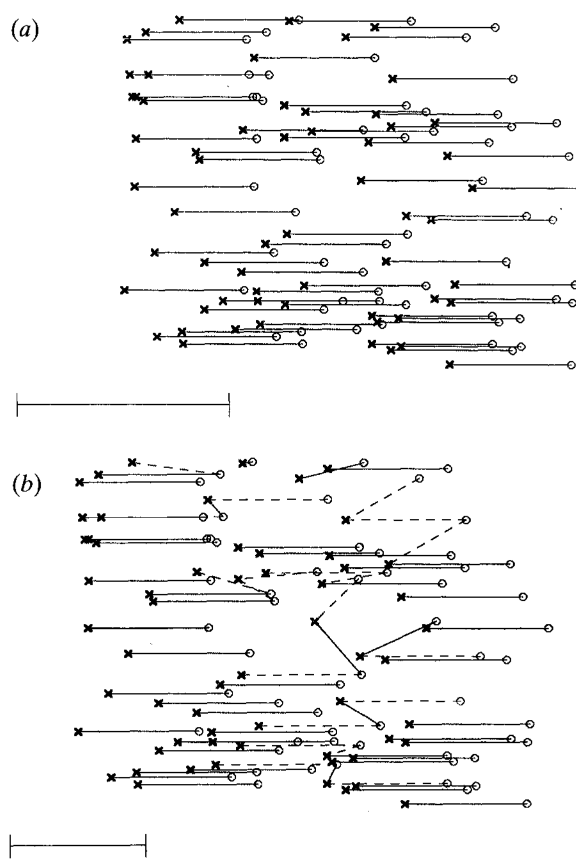

Figure 3. (*a*) Pure translation with no noise or 'rogue' points, and with a value of σ comparable to the displacement. (*b*) As (*a*), but with shorter-range σ. The flow is beginning to break up in the centre, but is not totally disrupted.

Figure 2*a*, *b* shows how the algorithm responds to the well-known 'Ternus configuration', given two slightly different values of σ. With σ equal to the separation between the images, P takes the form

0.814	−0.580
0.580	*0.814*

implying the 1:1 correspondence between crosses and circles illustrated in figure 2*a*; but with half this value of σ the principle of proximity defeats the principle of exclusion; the form of P

0.261	−0.965
0.965	0.261

implies that the right-hand cross is paired with the left-hand circle, the other two features remaining 'celibate'. There is a close analogy with movement perception. When the circles and the crosses are presented to human observers in quick succession, the perceived motion undergoes an analogous transition from a 'neighbour mode' to a 'coherent mode' as the inter-stimulus interval is increased (Ternus 1926; Ullman 1979). Ullman attributes such perceptual transitions to the 'affinities between basic elements' becoming more uniform in strength as the inter-stimulus interval is increased; this is exactly the effect of an increase in σ on the elements of the proximity matrix G.

Figures 3*a*, 4*a*, 5 and 6 show that with a sufficiently large value of σ the algorithm succeeds in recovering the 1:1 correspondences created by a translation, a shear deformation, an expansion and a combination of all three. Figure 3*b* shows the disruption to figure 3*a* that results from taking too small a value for σ. Figure 4*b* shows the effect of adding ten rogue points to the second image in figure 4*a* (the pattern of circles). Figure 7 shows that our algorithm is not very successful in correlating a pair of patterns related by a rotation; 'wagon-wheeling' starts to occur at some angle of rotation irrespective of the value of σ. Finally, figure 8 shows the correspondences that it finds between two frames from a real-world image sequence, when the features are 'corners' that have been identified in both frames. They are all veridical.

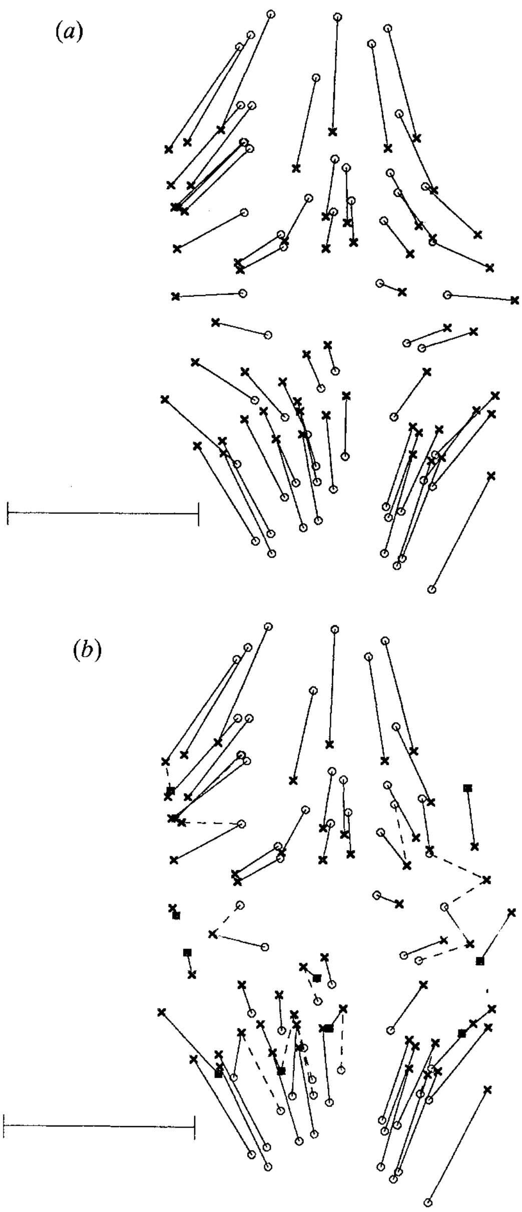

Figure 4. (*a*) The pattern of circles is generated from the (random) pattern of crosses by a shear deformation. (*b*) As figure 4*a*, except that ten rogue points (solid squares) have been added to the second image (circles).

5. IMAGES RELATED BY AFFINE TRANSFORMATIONS

In several of the examples we have described, and others too numerous to mention, the circles were derived from the crosses by an affine transformation not involving rotation, and in every case our algorithm (supplied with a sufficiently large value of σ), succeeds in finding the feature correspondences created by this transformation. Because successive images in a sequence will often be connected by transformations that are affine or nearly so, this property is one to be welcomed, if not positively required, in a satisfactory correspondence algorithm. The following argument is intended to explain why the algorithm performs so well in this respect.

The first stage in the argument is to show that if one set of points in a plane is mapped into another by a translation, an expansion or a shear deformation, then this 1:1 mapping minimizes the sum of the squares of the distances between corresponding points in the two sets. In the second stage we explain why, with a large enough value of σ, the mapping found by our algorithm possesses this particular property.

It may be helpful to introduce the first stage by considering the simple case in which the circles (at $\boldsymbol{s}_i$) are derived from the crosses (at $\boldsymbol{r}_i$) by a pure translation $\boldsymbol{t}$:

$$\boldsymbol{s}_i = \boldsymbol{r}_i + \boldsymbol{t}.$$

We proceed to show that any other 1:1 mapping results in a greater value for the sum

$$\Sigma(\boldsymbol{r}_i - \boldsymbol{s}_{i'})^2,$$

where i' denotes the new partner of the point i.

Because $(1', 2', \ldots)$ must be a permutation of $(1, 2, \ldots)$, and the general permutation is a product of independent cyclic permutations, we consider the

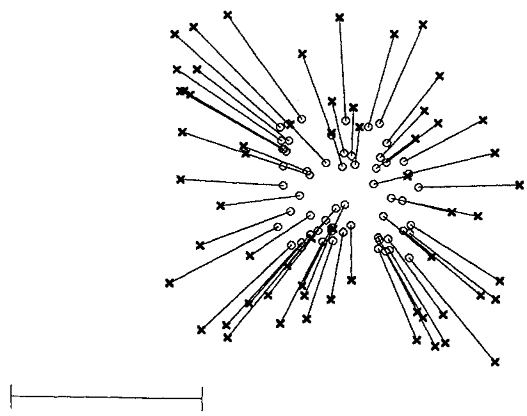

Figure 5. The pattern of circles is generated from the (random) pattern of crosses by a pure expansion.

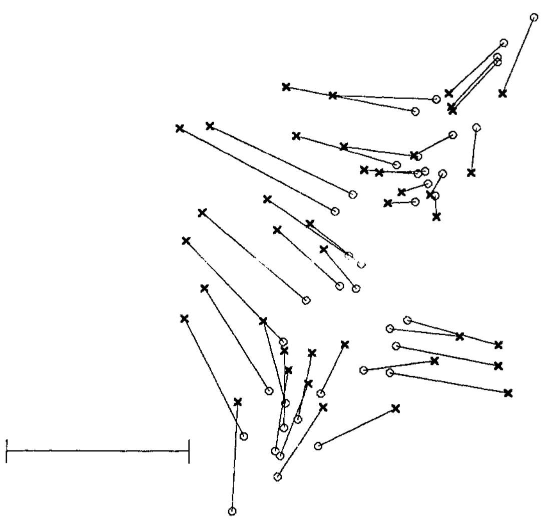

Figure 6. An arbitrary combination of translation, expansion and shear deformation. The algorithm discovers all the matches specified by this transformation, although this requires the crossing of arcs in several instances.

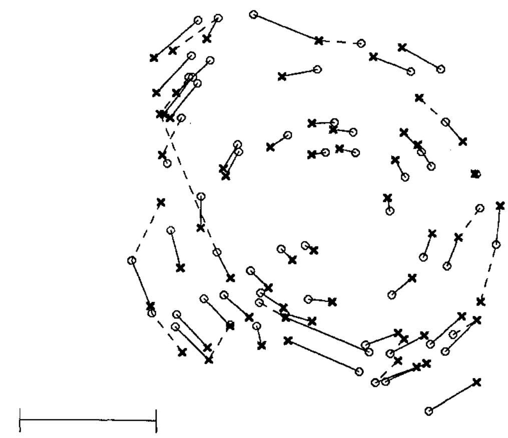

Figure 7. A random pattern of crosses subjected to a finite rotation. This is the most troublesome of the affine transformations, both for the algorithm and for people. 'Wagon-wheeling' starts to occur at a critical angle of rotation, regardless of the value of σ.

simplest non-trivial case $(1' = 2, 2' = 3, 3' = 1)$. The aim is then to show that

$$(\boldsymbol{r}_1-\boldsymbol{s}_2)^2+(\boldsymbol{r}_2-\boldsymbol{s}_3)^2+(\boldsymbol{r}_3-\boldsymbol{s}_1)^2 > (\boldsymbol{r}_1-\boldsymbol{s}_1)^2+(\boldsymbol{r}_2-\boldsymbol{s}_2)^2+(\boldsymbol{r}_3-\boldsymbol{s}_3)^2$$

i.e. that

$$2[\boldsymbol{r}_1(\boldsymbol{s}_1-\boldsymbol{s}_2)+\boldsymbol{r}_2(\boldsymbol{s}_2-\boldsymbol{s}_3)+\boldsymbol{r}_3(\boldsymbol{s}_3-\boldsymbol{s}_1)] > 0.$$

The latter inequality follows directly from the fact that $\boldsymbol{s}_1-\boldsymbol{s}_2 = \boldsymbol{r}_1-\boldsymbol{r}_2$, etc. and the fact that

$$(\boldsymbol{r}_1-\boldsymbol{r}_2)^2+(\boldsymbol{r}_2-\boldsymbol{r}_3)^2+(\boldsymbol{r}_3-\boldsymbol{r}_1)^2 > 0.$$

Next we turn to the general affine transformation

$$\boldsymbol{s}_i = \boldsymbol{A}\cdot\boldsymbol{r}_i+\boldsymbol{t},$$

in which $\boldsymbol{A}$ is a second-rank tensor and $\boldsymbol{t}$ (as before) is a translation vector. In Cartesian components this transformation takes the form

$$x'_i = ax_i+by_i+e, \quad y'_i = cx_i+dy_i+f,$$

and may be seen to comprise – loosely speaking – a translation (e,f), an expansion $(a+d)/2$, two components of shear, $(a-d)/2$ and $(b+c)/2$, and a rotation $(b-c)/2$. The first component of shear is mirror-symmetric about the x and y axes, and the other, of magnitude b, is antisymmetric about those axes. (These symmetries have nothing to do, of course, with the symmetry or lack of symmetry of $\boldsymbol{A}$ about its diagonal.)

At this point we restrict the tensor $\boldsymbol{A}$ to being both symmetric and positive definite, a property that ensures the positivity of the product $\boldsymbol{r}\cdot\boldsymbol{A}\cdot\boldsymbol{r}$ for every non-null vector $\boldsymbol{r}$. Necessary and sufficient conditions for this are that

$$a > 0, d > 0, ad > bc \quad \text{and} \quad b = c.$$

The last of these conditions precludes $\boldsymbol{A}$ from having any rotational component; the other conditions ensure that the transformation from circles to crosses never reverses the sense of a triangle of features. To complete the first stage in the argument we need to show that if the circles and the crosses are related by such a transformation, then any other pairing between them increases the sum of the squares of the distances between paired points. The algebra is much the same as before: one must show, for example, that

$$2[\boldsymbol{r}_1(\boldsymbol{s}_1-\boldsymbol{s}_2)+\boldsymbol{r}_2(\boldsymbol{s}_2-\boldsymbol{s}_3)+\boldsymbol{r}_3(\boldsymbol{s}_3-\boldsymbol{s}_1)] > 0.$$

Setting $\boldsymbol{s}_i = \boldsymbol{A}\cdot\boldsymbol{r}_i+\boldsymbol{t}$, and using the fact that $\boldsymbol{r}_i\cdot\boldsymbol{A}\cdot\boldsymbol{r}_j = \boldsymbol{r}_i\cdot\boldsymbol{A}\cdot\boldsymbol{r}_j$, we obtain the obviously positive expression

$$(\boldsymbol{r}_1-\boldsymbol{r}_2)\cdot\boldsymbol{A}\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)+(\boldsymbol{r}_2-\boldsymbol{r}_3)\cdot\boldsymbol{A}\cdot(\boldsymbol{r}_2-\boldsymbol{r}_3) + (\boldsymbol{r}_3-\boldsymbol{r}_1)\cdot\boldsymbol{A}\cdot(\boldsymbol{r}_3-\boldsymbol{r}_1),$$

and the above assertion follows.

The second stage in the argument is less rigorous. Returning to G, we note that if σ is large, then every element of G is close to 1, so that G is close to being a symmetric matrix. For such matrices, as shown earlier, the matrix P is the unit matrix; so $P = E$ is at least an approximate solution of the problem of maximizing $P{:}G$. But for a matrix G consisting entirely of 1s, $P{:}G$ is equally well maximized by any other matrix P having a single 1 in each row and in each column; so in this 'zeroth' approximation there is nothing to choose between E and these other permutation matrices, and we must go to the next approximation for further enlightenment. In this approximation we

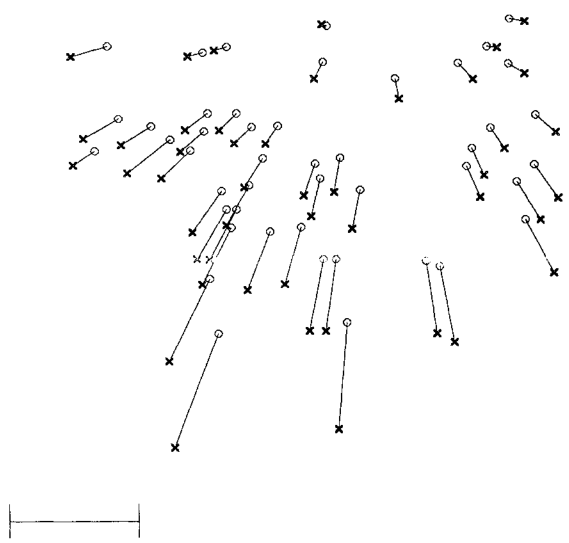

Figure 8. The performance of the algorithm as applied to two frames of a real-world image sequence. The correspondences are all veridical.

expand each Gaussian element in powers of $(1/\sigma^2)$, obtaining, to the first order in $(1/\sigma^2)$,

$$G_{ij} = 1 - r_{ij}^2/2\sigma^2.$$

The problem of finding the permutation matrix P that maximizes $P{:}G$ is now the problem of finding the 1:1 mapping from *i*s to *j*s that minimizes the sum of the squared distances r_{ij}^2 between the members of each pair. But if the positions of the circles and the crosses are indeed connected by an affine transformation with a positive definite tensor component, then the mapping between them does indeed minimize the sum of the squares of the distances. So at last we can see why, first of all, the P matrix delivered by our algorithm approximates to a permutation matrix; and secondly, why the permutation represented by the matrix faithfully recovers the 1:1 mapping originally induced by the affine transformation.

6. DISCUSSION

In its classical formulation the correspondence problem has to do with human vision, and the parameters that describe motion perception. But it is also a problem for computer vision engineers, and this is how we have chosen to address it. In the circumstances it is reassuring to find that our own solution to the problem has a certain similarity to Ullman's minimal mapping scheme, a theory that accounts for so many observations in the area of motion perception. There are, however, important differences.

Like us, Ullman aims to extremize the inner product of two matrices: one having to do with 'affinities' – or 'costs' – and the other a matrix of correspondences. He has much to say about such measures, though he does not explicitly consider the Gaussian proximity matrix, devoting much of his discussion to a cost measure that is linear in the separation. This leads him to minimize, where possible, the sum of the first powers of the distances between corresponding points, rather than the sum of their squares, which we find to be automatically minimized by a very general subset of the affine group.

Be that as it may, we attach more importance to the constraints to be placed on the form of the pairing matrix P. In our case the method of construction of P ensures its orthogonality; this tends to result in 1:1 mappings but does allow, in certain situations, what amount to 'splits' and 'fusions', without legislating for them in advance.

As for the relative merits of the two schemes, Ullman attaches considerable weight to the biological plausibility of his minimal mapping theory, while our algorithm, in its formal simplicity, is perhaps better suited to be a useful component of a computer vision system.

We thank Mike Brady for his encouragement and support, and Andrew Blake for useful discussions. H. C. L.-H. thanks the Leverhulme Foundation for financial support.

REFERENCES

Marr, D. 1976 Early processing of visual information. *Phil. Trans. Roy. Soc. Lond.* B **275**, 483–534.
Strang, G. 1988 *Linear algebra and its applications*, 3rd edn, Appendix A. San Diego: Harcourt Brace, Jovanovich.
Ternus, J. 1926 Experimentelle Untersuchung über phanomenale Identität. *Psychol. Forsch.* **7**, 81–136.
Ullman, S. 1979 *The interpretation of visual motion*, ch. 2 and 3. Cambridge, Massachusetts: M.I.T. Press.
Wertheimer, M. 1912 Experimentelle Studien uber das Sehen von Bewegung. *Z. Psychol.* **61**, 161–265.

APPENDIX 1

As stated above, the proximity matrix $G_{ij} = \exp(-r_{ij}^2/2\sigma^2)$ has the useful property of making G positive definite when its two subscripts run over the same n locations. To see why this is so, we consider for simplicity the one-dimensional case in which $\sigma = 1$ and $r_{ij} = |x_i - x_j|$. A necessary and sufficient condition for G to be positive definite is that for all real non-null vectors $(f_1, \ldots, f_n) = \boldsymbol{f}$, say, the quadratic form $\boldsymbol{f}G\boldsymbol{f}^T$ be positive. Given such a vector we introduce the distribution

$$f(x) = \Sigma_j f_j \delta(x - x_j),$$

where δ is the Dirac delta function, and its Fourier transform

$$g(k) = \sqrt{(1/2\pi)}\, \Sigma_j f_j \exp(-ikx_j),$$

and evaluate in two different ways the expression

$$(1/2\pi) \int dx \int dx' \int dk \int dk'\, g(k) \exp(ikx) \times \exp(-(x-x')^2/2) \exp(-ik'x')\, g^\dagger(k'),$$

in which each integral runs from $-\infty$ to $+\infty$. Integration over k and k' first, and then over x and x', gives the above-mentioned quadratic form; integration over x and x' first, and afterwards over k', yields the alternative expression

$$\int dk\, g(k) \exp(-k^2/2)\, g^\dagger(k),$$

which may be rewritten in the self-evidently positive form

$$\int |g(k) \exp(-k^2/4)|^2\, dk.$$

The extension to two dimensions (or more) is quite straightforward.

Received 17 December 1990; accepted 11 February 1991