Some experiments in fitting pairs of diagrams that lack defined reference points

P.H.A. Sneath
Department of Microbiology, Leicester University,
Leicester, LE1 7RH.

Abstract

The fitting of a pair of diagrams, and the measurement of their similarity in shape, is particularly difficult if the diagrams lack defined reference points. In such situations it is not possible to label corresponding points on two diagrams (e.g. point 1 on diagram A corresponds to point 1 on diagram B). The lack of correspondences may arise from doubts about homologies in biological structures, or difficulties in defining positions of interest on artefacts; further, choice of correspondences usually requires expert examination, and is not readily obtained by automated methods.

Some experiments will be described with a computer method that only requires the coordinates of m points of diagram A and n points of diagram B, provided these are not too few and are reasonably evenly-spaced. The diagrams need not be closed outlines, or of other special form. The method is based on a suggestion (Sneath, 1967, J. Zool., Lond. 151:68) that one measure of misfit is the sum of squared distances between each point in A and the nearest point in B, whichever point this is.

The diagrams are first overlapped at the centroids, scaled to unit two-dimensional variance to bring them to the same size, and then diagram B is rotated until the closest-point function mentioned above is a minimum. Diagram B is then reflected ("flipped") about the X axis, and the process repeated, so as to make provision for asymmetric objects. At angle of best fit the program also lists the closest points, g, h, i, ... on B to the points 1, 2, 3, ... on A. These can be viewed as "geometrically homologous" or "isologous" pairs of points. Fitting diagram A to diagram B gives non-identical but very similar results.

The curve of the distance function with angle is many-cusped, which leads to slow computation to avoid trapping in local minima. Attempts to smooth the curve (to obtain quicker search for the global minimum) will be described, together with results on various objects and possible extensions to more dimensions and to series of reconstructions that are intermediate in shape between the pairs of diagrams. The principle foreshadows (with digitizers and extremely fast computers) a very general and powerful method for pattern-recognition.

INTRODUCTION

There are a number of methods of finding the best fit between two diagrams on which have been marked pairs of corresponding reference points (Sneath, 1967; Bookstein, 1978, 1982). They also allow the measurement of similarity or misfit between the diagrams. The choice of reference points requires, however, decisions on homologies in the broad sense (not necessarily phylogenetic homology). Thus the bregma of the skull may be recognized as homologous in two specimens, but one cannot safely homologize the most posterior points on the skulls. Over most of the cranial vault there are no easily defined homologous points. Similar problems arise with artefacts: one might homologize the tips of two spearheads, but be unsure whether one should treat the broadest part of the spearheads as homologous.

The present contribution reports on some experiments with a method for matching a pair of diagrams on which no corresponding "homologous" points are available, e.g. the outline of two featureless artefacts. It is based on the strategy suggested in Sneath (1967), and leads both to the best fitting of the diagrams (with a measure of misfit) and the labelling of corresponding pairs of points. To avoid confusion with phylogenetic homology, the relationship between corresponding points is referred to as "geometric homology", but the term "isology" would also be appropriate (Florkin, 1962). It is assumed that in principle the diagrams are represented by digitized points that are reasonably closely and evenly spaced so as to represent the object in an adequate manner.

METHODS

The two diagrams, A and B are represented respectively by m and n points, whose coordinates on the X and Y axes are recorded. The two diagrams are first scaled so that the mean squared distance of the points from the centroid is unity, and they are overlapped at their centroids. This is achieved by replacing (for each diagram separately) the X, Y coordinates by x, y using the following transformation (Sneath, 1967):

$$x_i = (X_i - \bar{X})/s$$

$$y_i = (Y_i - \bar{Y}/s$$

where X_i, Y_i are the coordinates for point i. $\bar{X}$ and $\bar{Y}$ are the means of X and Y over m points (for A) or n points (for B), and

$$s = (s^2_X + s^2_Y)^{\frac{1}{2}}$$

where s^2_X and s^2_Y are the variances of X and Y (with m degrees of freedom for A and n degrees of freedom for B).

The remaining step is to rotate one diagram about the centroid until the misfit between the diagrams is a minimum. Because there are no corresponding marked pairs of points it was suggested (Sneath, 1967) that the rotation should be to that angle, g, that minimizes the sum of squared distances from each point of one diagram to the nearest point in the other. A BASIC program has been written for this, and it can be run for moderate numbers of points on a microcomputer.

The minimum distance function is defined, for diagram A held fixed, and B rotated, as follows (it is convenient to employ the mean, so that it is independent of m):

$$\bar{d}^2_a = \frac{1}{m} \sum_{i=1}^{m} d^2_{ic}$$

where d^2_{ic} is the squared distance of point i in A to the closest point, c, in B, i.e. $(x_i - x_c)^2 + (y_i - y_c)^2$. Point c may be the same for several points in A; conversely, a particular point in B may never become c (i.e. it may never be a "closest point").

Three points should be noted: (1) the fitting of B to A is not exactly the same as that of A to B; (2) with asymmetric objects it is necessary to reflect one diagram about an axis (to "flip it over"), and determine both configurations; and (3) when a diagram is rotated the curve of $\bar{d}^2_a$ shows a number of local

minima. The first is self-evident: the number of points in A and B may differ, so one expects some discrepancy in the exact values for $\underline{d}^2_{\underline{a}}$. The second point is readily explained if one considers two coins, which might show identical portraits but facing left in one, right in the other. One cannot superimpose the portraits unless one is reflected from left to right. The third point is most obviously illustrated by radially-symmetrical objects. A pair of five-limbed starfish will match well at five positions where the limbs are coincident. If some additional feature were shown, e.g. the madreporite (which is eccentric), there will be a global minimum, and this is the desired solution. These three factors make the fitting very slow,— particularly the third, because it is necessary to search at close angles in order to find the global minimum.

RESULTS

The analysis of objects of simple outline and bilateral symmetry is shown in Figs. 1-7.

Fig. 1 shows the outlines of two handaxes, with the original coordinates in 1a and the coordinates after scaling and overlapping at the centroids in 1b. They show the minimum number of points needed to represent the outlines reasonably well. Note that no information is provided on homologous points, e.g. the tip is numbered 7 in A but 1 in B.

When B is rotated, and the sum of $\underline{d}^2_{\underline{a}}$ determined at eight equally-spaced angles, one observes (Fig. 2) two troughs (where the handaxes lie side by side) and two peaks (where they lie athwart one another), and this is seen both for the "unflipped" case (Fig. 2a) and the "flipped" case (Fig. 2b) in which B was reflected before rotation about the $\underline{X}$ axis. When the configuration was searched close to the four minima, the results are shown in Fig. 3a and 3b, which show the best fitting of B to diagram A. In 3a, diagram B is shown in the best, i.e. global, fit for the "unflipped" case by dashed lines, and for the "flipped" case by the dotted lines. In Fig. 3b, the minor minima are similarly shown. The figures also show the "geometric homologies" in the form of a vector of numbers that list the closest points in B to the eight points in A. Thus in 3a on the left the vector 66112344 indicates that the points 1, 2, 3, 4, 5, 6, 7, 8 in A

have as "geometrically homologous" points in B the points 6, 6, 1, 1, 2, 3, 4, 4 respectively.

In Figs. 3c and 3d are shown the corresponding results for diagram A rotated and fitted to B. Note that the vectors of "geometrical homology" have here only six members, although there are eight in A.

The values for $\underline{d}^2_{\underline{a}}$ and rotation angle in degrees are also shown in Fig. 3, those for the "unflipped" case on the left, and the "flipped" case on the right. It can be seen that the misfit values (mean $\underline{d}^2$) are very similar for all the minima, whether the global or the minor minima. This is to be expected, because the handaxes are not only roughly bilaterally symmetric, but are not very obviously pointed at one end rather than the other,– at least not when represented by so few points.

The problem of finding the global minimum is illustrated by Fig. 4, which shows the curve of the sum of $\underline{d}^2_{\underline{a}}$ for B fitted to A "unflipped", for 72 angles at 5^o intervals. There are seven visible minima, but one requires the global minimum near 45^o. Fig. 5 shows the region near this enlarged, with closer spaced angles, and the inset shows yet finer detail. It is evident there are a series of cusps, one of which is clearly shown, and these correspond to the points at which the "geometrical homology" vectors change (shown by vertical bars in Figs. 4 and 5).

These cusps pose exceptional difficulties for iterative searches. The computing method therefore employs an exhaustive search, analogous to the well-known method of repeated bisection, by determining $\Sigma\underline{d}^2_{\underline{a}}$ at $\underline{r}$ equally-spaced intervals and reducing at each cycle the range represented by these $\underline{r}$ angles, (centering them successively on the previous angle with lowest $\Sigma\underline{d}^2_{\underline{a}}$). The time required is almost independent of $\underline{r}$, so it is wise to make $\underline{r}$ fairly large (e.g. 32) so as to search the entire circle of 360^o closely during the first cycle.

The global minimum of mean $\underline{d}^2_{\underline{a}}$ for unflipped or flipped configuration, can then be taken as a measure of shape dissimilarity $\underline{d}^2_{\underline{ag}}$. When B has been fitted to A the process can be repeated by fitting diagram A to B. Results, though not identical as has been noted, are commonly very close to those for fitting B to A.

The points in Figs. 1-5 were not chosen to be numerous or accurately representative of the outlines. When the numbers of

points are increased, the fitting is obviously improved (Figs. 6, 7), and it is clear that the pointed ends of the handaxes now have a greater influence on the fitting process. The cusps, too, tend to be less marked, though they are still present and are more numerous. In Figs. 6 and 7 the "unflipped" and "flipped" solutions are still very close, because the handaxes are almost bilaterally symmetrical: the difference between the solution could be used as a measure of asymmetry. It can be seen that the fit of B to A (Fig. 7a) is almost identical to that of A to B (Fig. 7b).

The method is not restricted to closed outlines as in Fig. 1, but can handle arbitary shapes. Figs. 8 and 9 show the fitting of capital letters E and F. The fitting of F to E is not a trivial problem when it is remembered that no information is given to the computer to the effect that F is an alphabetic letter consisting of E without a lower cross-bar.

In an attempt to speed up the computation an alternative search was also explored. In this, the quadratic correlation of the lowest five values of $\Sigma \underline{d}^2_{\underline{a}}$ was determined for each iteration, and when this became small a quadratic function was fitted, and searched for the minimum. This was not very successful, and sometimes led to trapping in a local minimum.

More successful was a harmonic function. The rationale is that abrupt changes and cusps could be avoided if the distances from point $\underline{i}$ on diagram A to all points on B were considered, not only the single distance to the nearest point, $\underline{c}$. The distances must be weighted inversely to magnitude, but they must be protected from a reciprocal of zero if two points should become superimposed. This can be done by adding to all distances a small quantity, chosen to reflect approximately the average distance to be expected between points. The harmonic function for B fitted to A was

$$\underline{d}^2_{\underline{c}} = \underline{m} \Big/ \left[\sum_{\underline{h}=1}^{\underline{m}} \sum_{\underline{i}=1}^{\underline{n}} \left(1/(\underline{d}^2_{\underline{hi}} + (\underline{k}/\underline{m})) \right) \right]$$

where $\underline{h}$ and $\underline{i}$ are points on A and B respectively and $\underline{d}^2_{\underline{hi}}$ is $(\underline{x}_{\underline{h}} - \underline{x}_{\underline{i}})^2 + (\underline{y}_{\underline{h}} - \underline{y}_{\underline{i}})^2$. A suitable value for $\underline{k}$ is unity.

The curve of $\underline{d}^2_{\underline{c}}$ against angle was much smoother than that of $\underline{d}^2_{\underline{a}}$, and permitted easier quadratic search, with considerable improvement in speed. The angle of best fit from $\underline{d}^2_{\underline{c}}$ is not quite the same as for $\underline{d}^2_{\underline{a}}$ but was usually within 0.5°.

have as "geometrically homologous" points in B the points 6, 6, 1, 1, 2, 3, 4, 4 respectively.

In Figs. 3c and 3d are shown the corresponding results for diagram A rotated and fitted to B. Note that the vectors of "geometrical homology" have here only six members, although there are eight in A.

The values for $\underline{d}^2_{\underline{a}}$ and rotation angle in degrees are also shown in Fig. 3, those for the "unflipped" case on the left, and the "flipped" case on the right. It can be seen that the misfit values (mean $\underline{d}^2$) are very similar for all the minima, whether the global or the minor minima. This is to be expected, because the handaxes are not only roughly bilaterally symmetric, but are not very obviously pointed at one end rather than the other,— at least not when represented by so few points.

The problem of finding the global minimum is illustrated by Fig. 4, which shows the curve of the sum of $\underline{d}^2_{\underline{a}}$ for B fitted to A "unflipped", for 72 angles at 5° intervals. There are seven visible minima, but one requires the global minimum near 45°. Fig. 5 shows the region near this enlarged, with closer spaced angles, and the inset shows yet finer detail. It is evident there are a series of cusps, one of which is clearly shown, and these correspond to the points at which the "geometrical homology" vectors change (shown by vertical bars in Figs. 4 and 5).

These cusps pose exceptional difficulties for iterative searches. The computing method therefore employs an exhaustive search, analogous to the well-known method of repeated bisection, by determining $\Sigma\underline{d}^2_{\underline{a}}$ at $\underline{r}$ equally-spaced intervals and reducing at each cycle the range represented by these $\underline{r}$ angles, (centering them successively on the previous angle with lowest $\Sigma\underline{d}^2_{\underline{a}}$). The time required is almost independent of $\underline{r}$, so it is wise to make $\underline{r}$ fairly large (e.g. 32) so as to search the entire circle of 360° closely during the first cycle.

The global minimum of mean $\underline{d}^2_{\underline{a}}$ for unflipped or flipped configuration, can then be taken as a measure of shape dissimilarity $\underline{d}^2_{\underline{ag}}$. When B has been fitted to A the process can be repeated by fitting diagram A to B. Results, though not identical as has been noted, are commonly very close to those for fitting B to A.

The points in Figs. 1-5 were not chosen to be numerous or accurately representative of the outlines. When the numbers of

points are increased, the fitting is obviously improved (Figs. 6, 7), and it is clear that the pointed ends of the handaxes now have a greater influence on the fitting process. The cusps, too, tend to be less marked, though they are still present and are more numerous. In Figs. 6 and 7 the "unflipped" and "flipped" solutions are still very close, because the handaxes are almost bilaterally symmetrical: the difference between the solution could be used as a measure of asymmetry. It can be seen that the fit of B to A (Fig. 7a) is almost identical to that of A to B (Fig. 7b).

The method is not restricted to closed outlines as in Fig. 1, but can handle arbitary shapes. Figs. 8 and 9 show the fitting of capital letters E and F. The fitting of F to E is not a trivial problem when it is remembered that no information is given to the computer to the effect that F is an alphabetic letter consisting of E without a lower cross-bar.

In an attempt to speed up the computation an alternative search was also explored. In this, the quadratic correlation of the lowest five values of Σd^2_a was determined for each iteration, and when this became small a quadratic function was fitted, and searched for the minimum. This was not very successful, and sometimes led to trapping in a local minimum.

More successful was a harmonic function. The rationale is that abrupt changes and cusps could be avoided if the distances from point i on diagram A to all points on B were considered, not only the single distance to the nearest point, c. The distances must be weighted inversely to magnitude, but they must be protected from a reciprocal of zero if two points should become superimposed. This can be done by adding to all distances a small quantity, chosen to reflect approximately the average distance to be expected between points. The harmonic function for B fitted to A was

$$d^2_c = m \Big/ \left[\sum_{h=1}^{m} \sum_{i=1}^{n} \left(1/(d^2_{hi} + (k/m)) \right) \right]$$

where h and i are points on A and B respectively and d^2_{hi} is $(x_h - x_i)^2 + (y_h - y_i)^2$. A suitable value for k is unity.

The curve of d^2_c against angle was much smoother than that of d^2_a, and permitted easier quadratic search, with considerable improvement in speed. The angle of best fit from d^2_c is not quite the same as for d^2_a but was usually within 0.5^o.

have as "geometrically homologous" points in B the points 6, 6, 1, 1, 2, 3, 4, 4 respectively.

In Figs. 3c and 3d are shown the corresponding results for diagram A rotated and fitted to B. Note that the vectors of "geometrical homology" have here only six members, although there are eight in A.

The values for $\underline{d}^2_a$ and rotation angle in degrees are also shown in Fig. 3, those for the "unflipped" case on the left, and the "flipped" case on the right. It can be seen that the misfit values (mean $\underline{d}^2$) are very similar for all the minima, whether the global or the minor minima. This is to be expected, because the handaxes are not only roughly bilaterally symmetric, but are not very obviously pointed at one end rather than the other,— at least not when represented by so few points.

The problem of finding the global minimum is illustrated by Fig. 4, which shows the curve of the sum of $\underline{d}^2_a$ for B fitted to A "unflipped", for 72 angles at 5^o intervals. There are seven visible minima, but one requires the global minimum near 45^o. Fig. 5 shows the region near this enlarged, with closer spaced angles, and the inset shows yet finer detail. It is evident there are a series of cusps, one of which is clearly shown, and these correspond to the points at which the "geometrical homology" vectors change (shown by vertical bars in Figs. 4 and 5).

These cusps pose exceptional difficulties for iterative searches. The computing method therefore employs an exhaustive search, analogous to the well-known method of repeated bisection, by determining $\Sigma\underline{d}^2_a$ at $\underline{r}$ equally-spaced intervals and reducing at each cycle the range represented by these $\underline{r}$ angles, (centering them successively on the previous angle with lowest $\Sigma\underline{d}^2_a$). The time required is almost independent of $\underline{r}$, so it is wise to make $\underline{r}$ fairly large (e.g. 32) so as to search the entire circle of 360^o closely during the first cycle.

The global minimum of mean $\underline{d}^2_a$ for unflipped or flipped configuration, can then be taken as a measure of shape dissimilarity $\underline{d}^2_{ag}$. When B has been fitted to A the process can be repeated by fitting diagram A to B. Results, though not identical as has been noted, are commonly very close to those for fitting B to A.

The points in Figs. 1-5 were not chosen to be numerous or accurately representative of the outlines. When the numbers of

points are increased, the fitting is obviously improved (Figs. 6, 7), and it is clear that the pointed ends of the handaxes now have a greater influence on the fitting process. The cusps, too, tend to be less marked, though they are still present and are more numerous. In Figs. 6 and 7 the "unflipped" and "flipped" solutions are still very close, because the handaxes are almost bilaterally symmetrical: the difference between the solution could be used as a measure of asymmetry. It can be seen that the fit of B to A (Fig. 7a) is almost identical to that of A to B (Fig. 7b).

The method is not restricted to closed outlines as in Fig. 1, but can handle arbitary shapes. Figs. 8 and 9 show the fitting of capital letters E and F. The fitting of F to E is not a trivial problem when it is remembered that no information is given to the computer to the effect that F is an alphabetic letter consisting of E without a lower cross-bar.

In an attempt to speed up the computation an alternative search was also explored. In this, the quadratic correlation of the lowest five values of $\Sigma \underline{d}^2_{\underline{a}}$ was determined for each iteration, and when this became small a quadratic function was fitted, and searched for the minimum. This was not very successful, and sometimes led to trapping in a local minimum.

More successful was a harmonic function. The rationale is that abrupt changes and cusps could be avoided if the distances from point $\underline{i}$ on diagram A to all points on B were considered, not only the single distance to the nearest point, $\underline{c}$. The distances must be weighted inversely to magnitude, but they must be protected from a reciprocal of zero if two points should become superimposed. This can be done by adding to all distances a small quantity, chosen to reflect approximately the average distance to be expected between points. The harmonic function for B fitted to A was

$$\underline{d}^2_{\underline{c}} = \underline{m} \Big/ \left[\sum_{\underline{h}=1}^{\underline{m}} \sum_{\underline{i}=1}^{\underline{n}} \left(1/(\underline{d}^2_{\underline{hi}} + (\underline{k}/\underline{m})) \right) \right]$$

where $\underline{h}$ and $\underline{i}$ are points on A and B respectively and $\underline{d}^2_{\underline{hi}}$ is $(\underline{x}_{\underline{h}} - \underline{x}_{\underline{i}})^2 + (\underline{y}_{\underline{h}} - \underline{y}_{\underline{i}})^2$. A suitable value for $\underline{k}$ is unity.

The curve of $\underline{d}^2_{\underline{c}}$ against angle was much smoother than that of $\underline{d}^2_{\underline{a}}$, and permitted easier quadratic search, with considerable improvement in speed. The angle of best fit from $\underline{d}^2_{\underline{c}}$ is not quite the same as for $\underline{d}^2_{\underline{a}}$ but was usually within 0.5°.

DISCUSSION

The harmonic function is not itself very attractive as a misfit measure, because $\underline{d}^2_{\underline{c}}$ is not zero for identical diagrams and its algebraic properties are difficult to derive. But it may be useful to find the angle of global fit, and then employ $\underline{d}^2_{\underline{a}}$ as a misfit value. For identical diagrams $\underline{d}^2_{\underline{a}}$ is zero, but it should be noted that it can be zero with non-identical diagrams for either A fitted to B or B fitted to A: if A and B are identical except that B has an extra outlying point, $\underline{h}$, then $\underline{d}^2_{\underline{a}}$ of A to B may be zero because $\underline{h}$ is never a closest point, $\underline{c}$; for B to A, however, $\underline{h}$ will make a non-zero contribution toward $\underline{d}^2_{\underline{a}}$.

The method is obviously sensitive to large differences in shape, particularly major regional changes, as when one region of A is greatly compressed in B. The fitting is an overall fit for the diagrams as wholes, and no provision is made for major local distortion. Therefore the "geometric homologies" can be inappropriate in some regions of the diagrams. One possible solution is to construct a new diagram I intermediate between A and B, represented by the midpoints of the lines joining "geometric homologies", and to fit A to I and I to B.

Imagine other intermediates, similarly constructed, between I and B, and extend this process indefinitely. For convenience replace A by 0 and B by 1; then I is symbolized by 1/2. One can then visualize a chain of intermediates, as finely spaced as desired, e.g. 0, 1/8, 1/4, 3/8, 1/2, ... 7/8, 1. It may be that the chain of "geometric homology" vectors between successive items of the series would offer a way of improving on the imperfect homologization of the two ends of the series. Siegel and Benson (1982) describe a median fitting that may be applicable here. Rotational fit methods can be applied to three dimensions (Sneath, 1967; Perkins and Green, 1982) and can be extended to many dimensions (Gower, 1971). The present method can similarly be extended if required.

The method is a first attempt at a complex problem, and clearly needs to be improved. It is of interest because of its generality. Pattern recognition methods are commonly restricted to certain shapes. Rohlf and Ferson (1983) in reviewing the biological applications note that even with relatively simple shapes there are considerable problems in automated image analysis (Vanderheydt et al., 1979). Simple closed outlines can be handled by Fourier

analysis (e.g. Kaesler and Waters, 1972; Ferson et al., 1985), and principal components is sometimes applicable (Snee and Andrews, 1971; Wold, 1976). Moment invariants (Hu, 1962) are subject to serious drawbacks, as they often lead to intractable indeterminacies (Rohlf, personal communication). The present method can be adapted to closely spaced points obtained from automatic digitizers, and one can envisage that with high-speed computers it could be a powerful and general method for pattern recognition.

REFERENCES

Bookstein, F.L. 1978. The measurement of biological shape and shape change. Lecture notes in biomathematics, No. 24. Springer Verlag, New York.

Bookstein, F.L. 1982. Foundations of morphometrics. Ann. Rev. Ecol. Syst. 13: 451-470.

Ferson, S., Rohlf, F.J. and Koehn, R.K. 1985. Measuring shape variation of two-dimensional outlines. Syst. Zool. 34: 59-68.

Florkin, M. (1962). Isologie, homologie, analogie et convergence en biochimie comparée. Bull. Acad. Roy. Belg. Classe de Sci., Sér 5 48: 819-824.

Gower, J.C. 1971. Statistical methods of comparing different multivariate analyses of the same data. Pp. 138-149 in Mathematics in the archaeological and historical sciences. F.R. Hodson, D.G. Kendall and P. Tăutu, eds. Edinburgh Univ. Press, Edinburgh.

Hu, M.K. 1962. Visual pattern recognition by moment invariants IRE Trans. Information Theory 8: 179-187.

Kaesler, R.L. and Waters, J.A. 1972. Fourier analysis of the ostracod margin. Bull. Geol. Soc. Amer. 83: 1169-1178.

Perkins, W.J. and Green, R.J. 1982. Three-dimensional reconstruction of biological sections. J. Biomed. Engin. 4: 37-43.

Rohlf, F.J. and Ferson, S. 1983. Image analysis. Pp. 583-599 in Numerical Taxonomy, J. Felsenstein, ed. NATO ASI Series G, Ecological Sciences, No. 1. Springer Verlag, New York.

Siegal, A.F. and Benson, R.H. 1982. A robust comparison of biological shapes. Biometrics 38: 341-350.

Sneath, P.H.A. 1967. Trend-surface analysis of transformation grids. J. Zool. Lond. 151: 65-122.

Snee, R.D. and Andrews, H.P. 1971. Statistical design and analysis of shape studies. Appl. Statist. 20: 250.

Vanderheydt, L., Oosterlinck, A. and VanderBerghe, H. 1979. Design of a special interpreter for the classification of human chromosomes. IEEE Trans. Pattern Recog. Mach. Intell. PAMI-1: 214-219.

Wold, S. 1976. Pattern recognition by means of disjoint principal component models. Pattern recognition 8: 127-138.

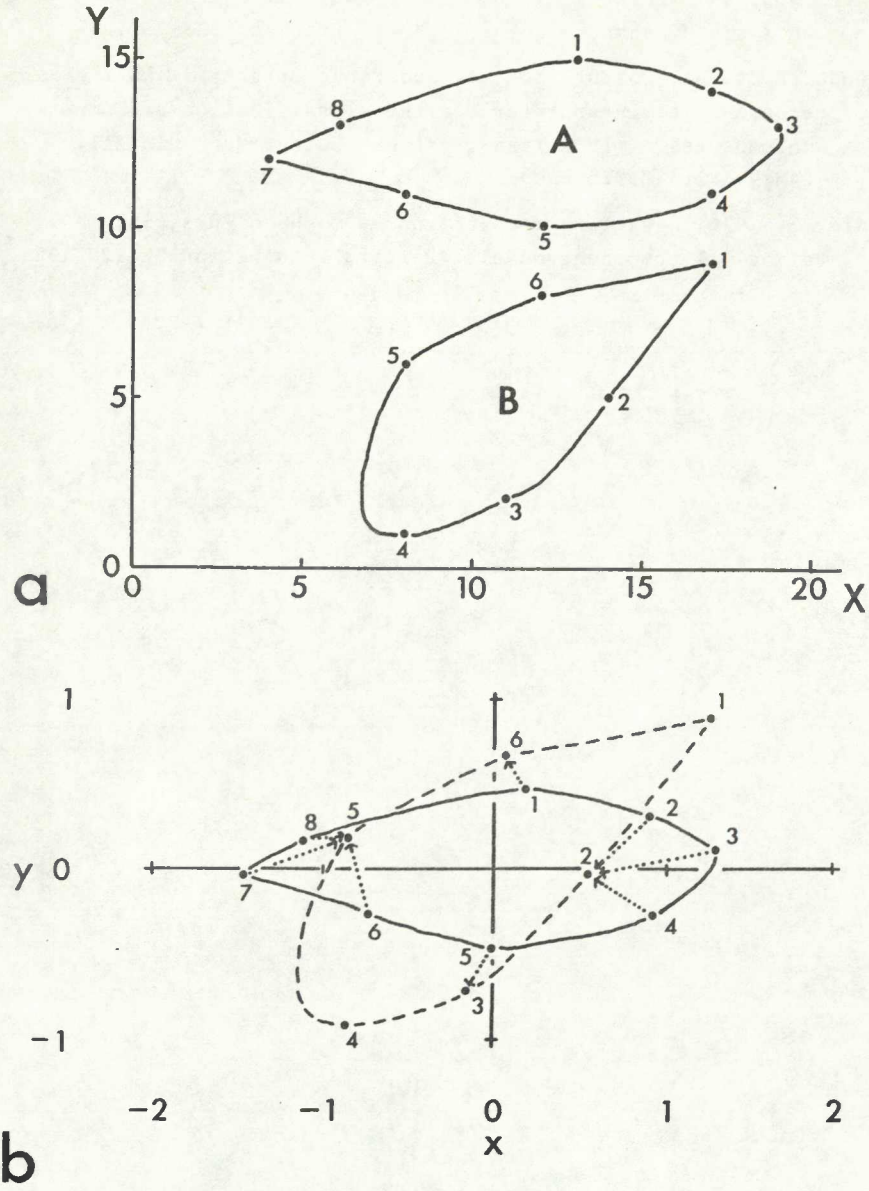

Fig.1. Two handaxes with 8 points marked in A and 6 points in B: (a) original co-ordinates; (b) coordinates after overlapping and scaling.

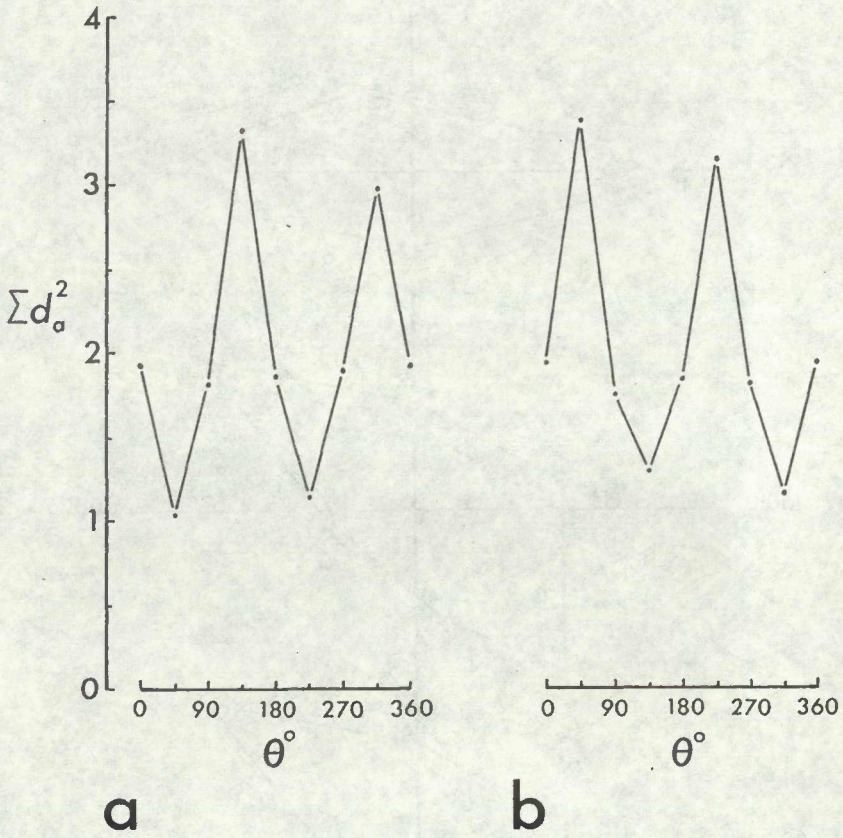

Fig.2. Values of the sum of d^2_a for rotation of B (Fig. 1b) at eight angles: (a) "unflipped" configuration; (b) "flipped" configuration.

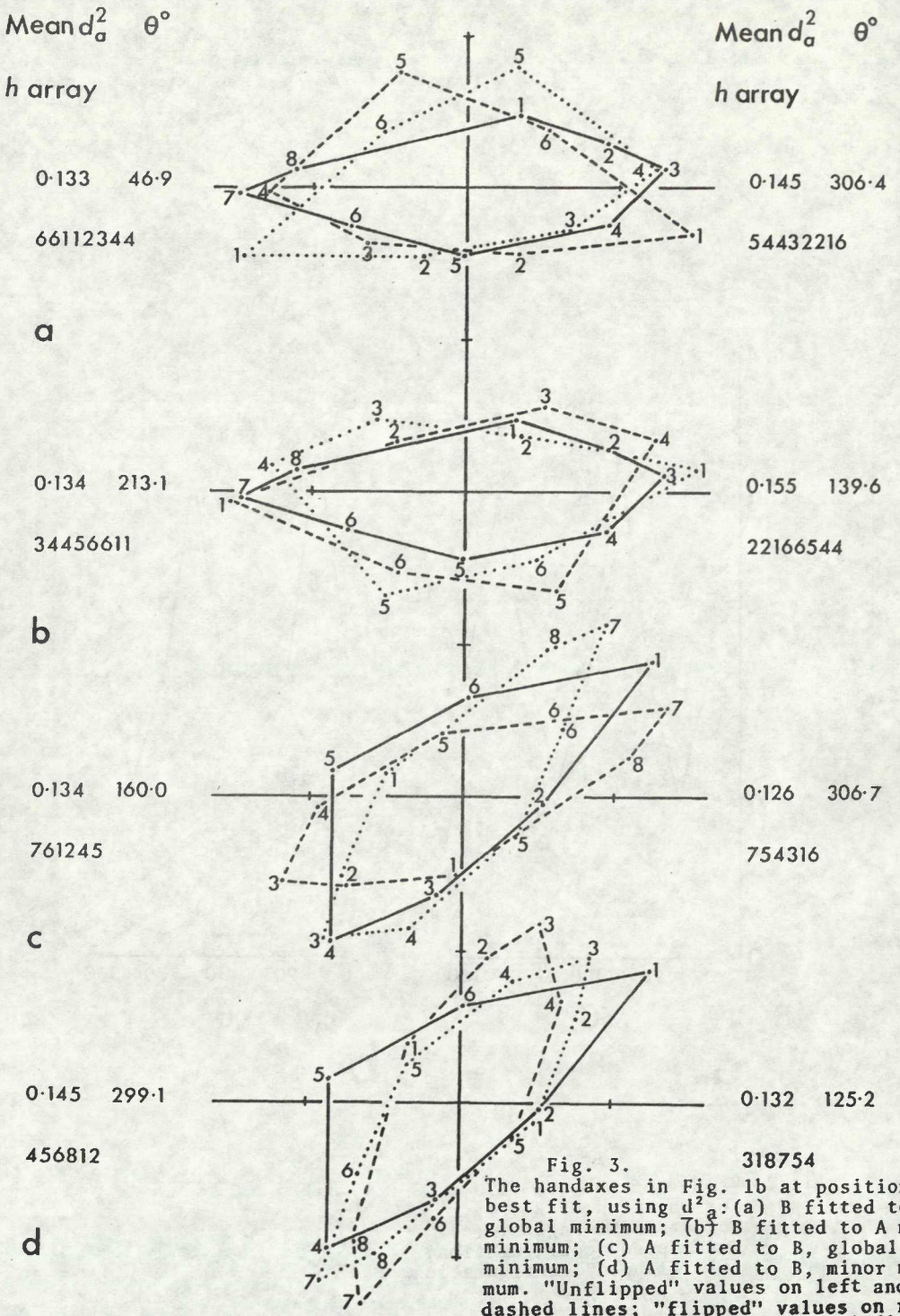

Fig. 3.
The handaxes in Fig. 1b at position of best fit, using d^2_a: (a) B fitted to A, global minimum; (b) B fitted to A minor minimum; (c) A fitted to B, global minimum; (d) A fitted to B, minor minimum. "Unflipped" values on left and dashed lines; "flipped" values on right and dotted lines. Solid lines indicate unrotated diagram.

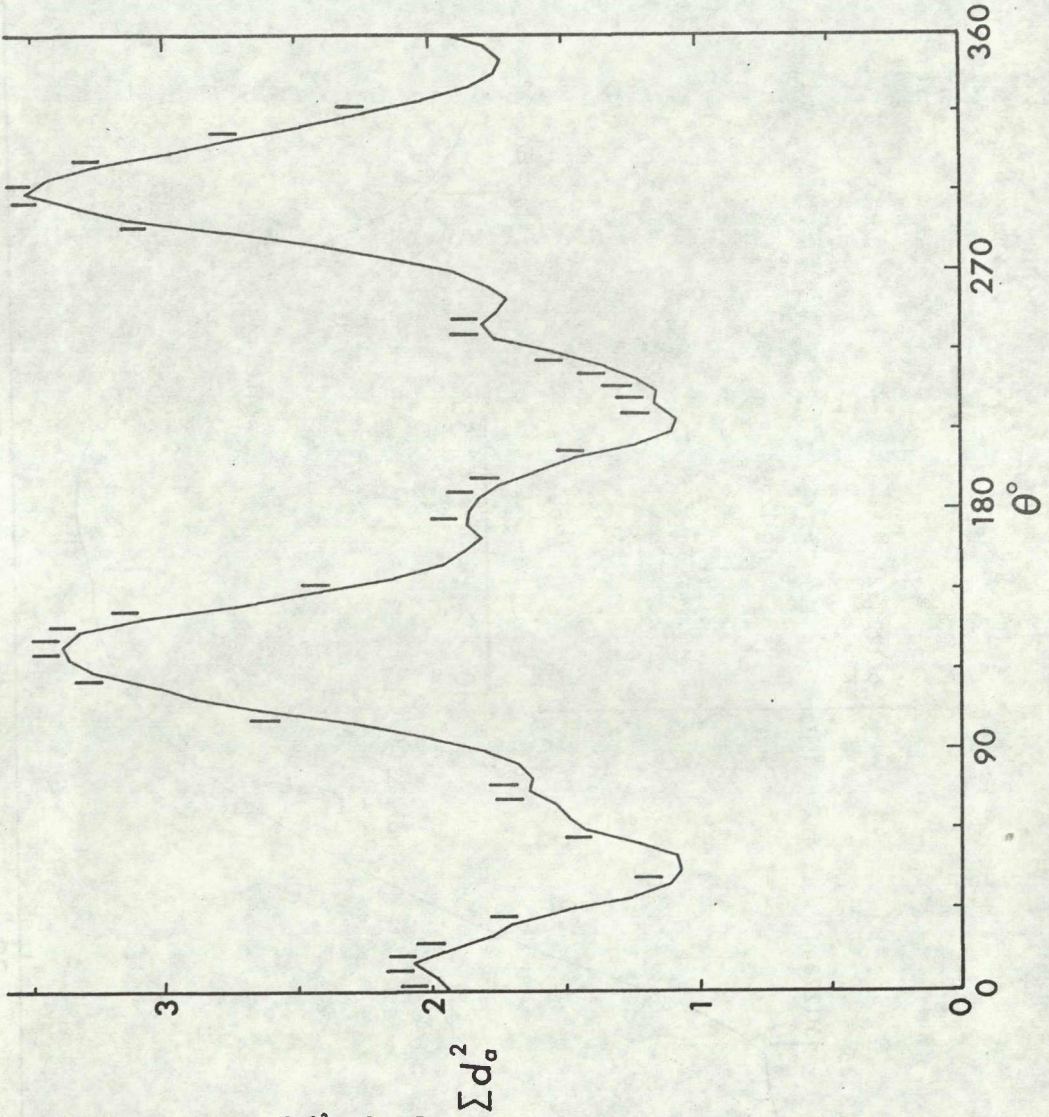

Fig.4. Values of sum of $\underline{d}^2_{\underline{a}}$ for B fitted to A (Fig. 1$\bar{b}$) at 5^o intervals. Vertical bars indicate where vectors of "geometric homology" change.

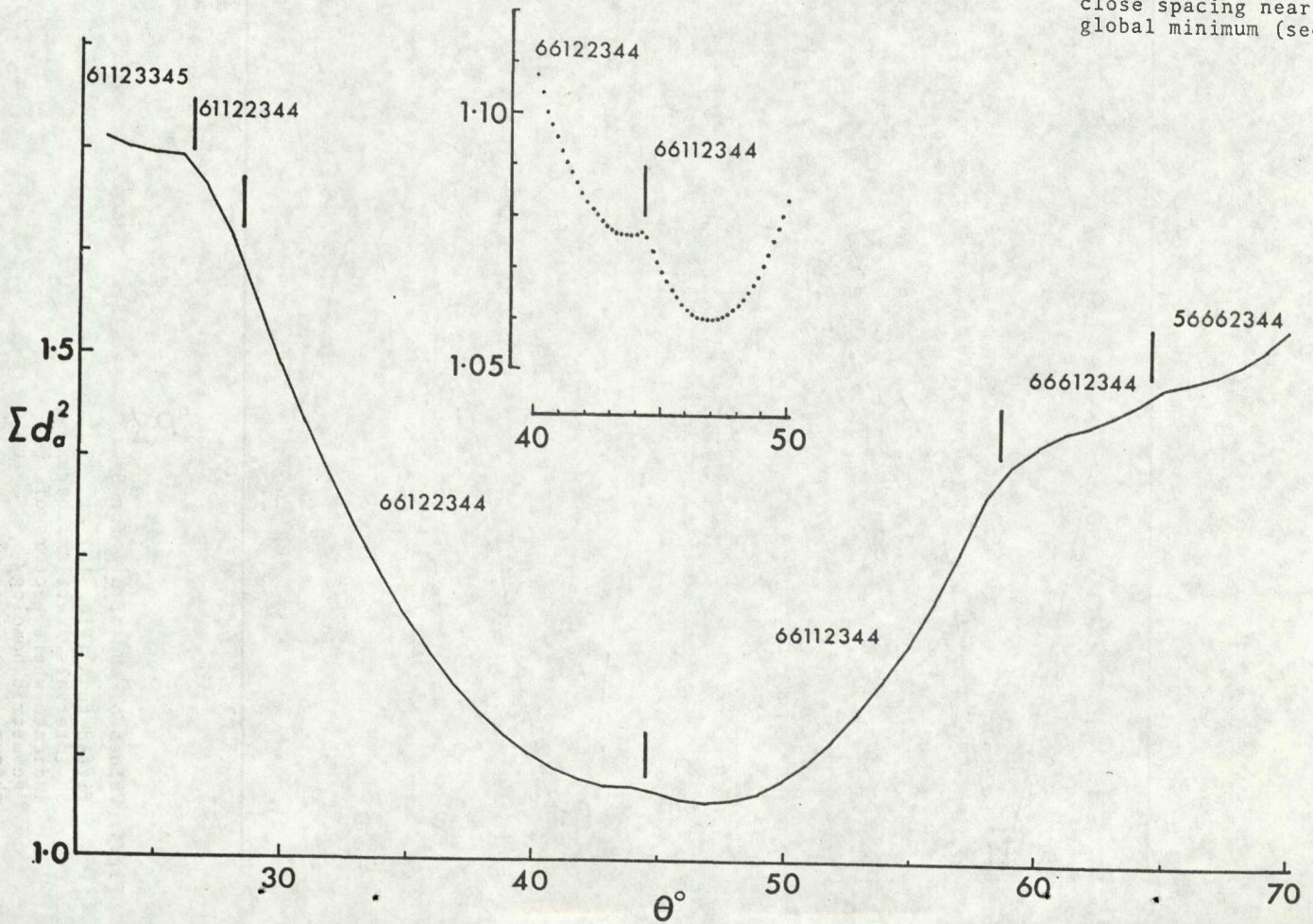

Fig.5. Values of sum of $\underline{d}^2_a$ for B fitted to A (Fig. 1$\bar{b}$) at close spacing near the global minimum (see Fig. 4).

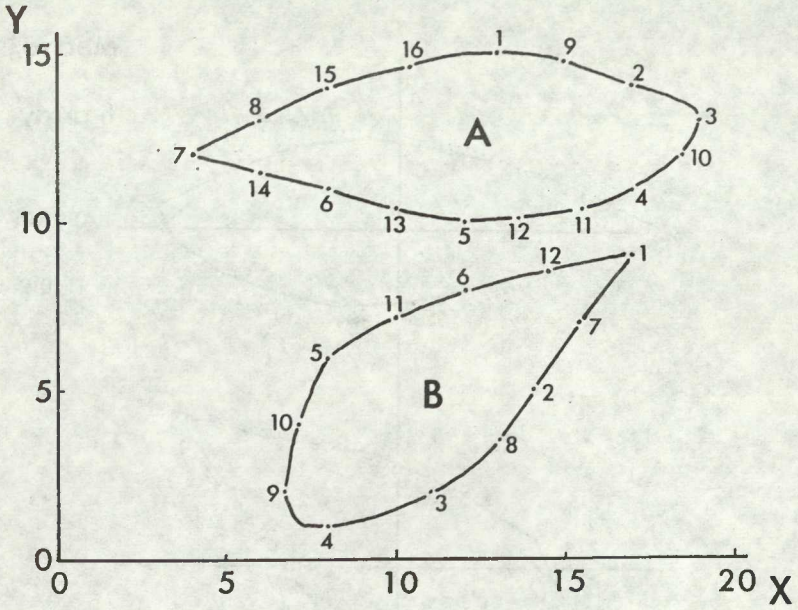

Fig.6. Handaxes of Fig. 1 represented by more closely-spaced points, 16 in A and 12 in B.

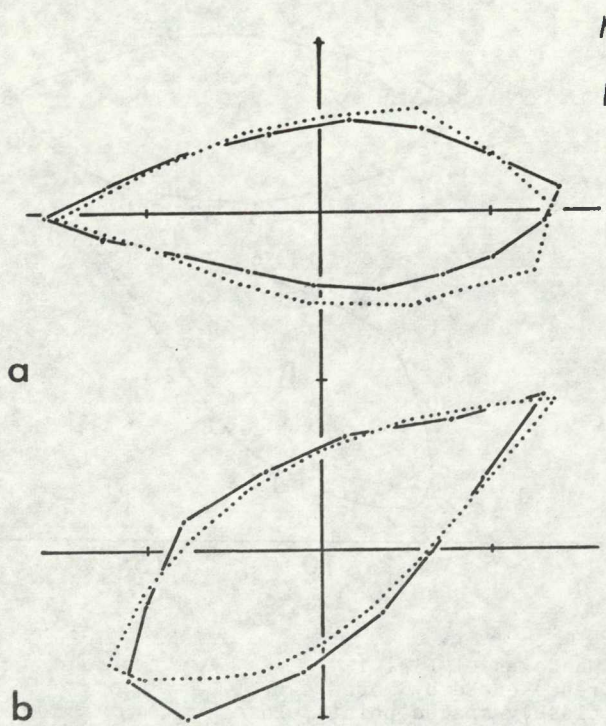

Fig.7. Handaxes of Fig. 6 at best global fit: (a) B fitted to A; (b) A fitted to B; solid lines indicate unrotated diagram, dotted lines rotated diagram. Other conventions as in Fig. 3.

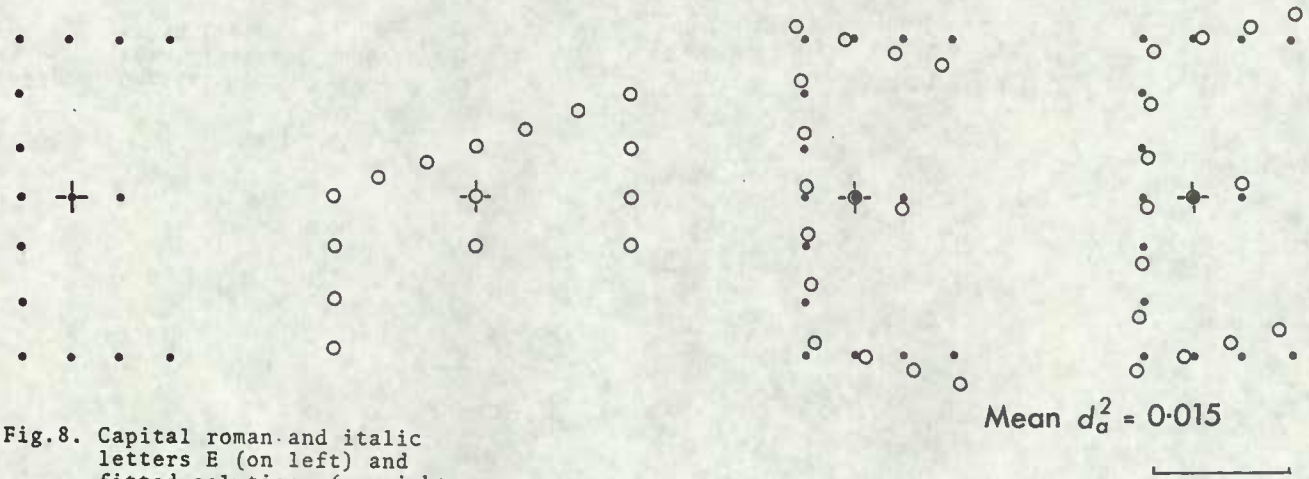

Fig.8. Capital roman and italic letters E (on left) and fitted solutions (on right; "flipped" and "unflipped" solutions give equal minima in $\underline{d}^2\underline{a}$. The bar indicates one standard deviation.

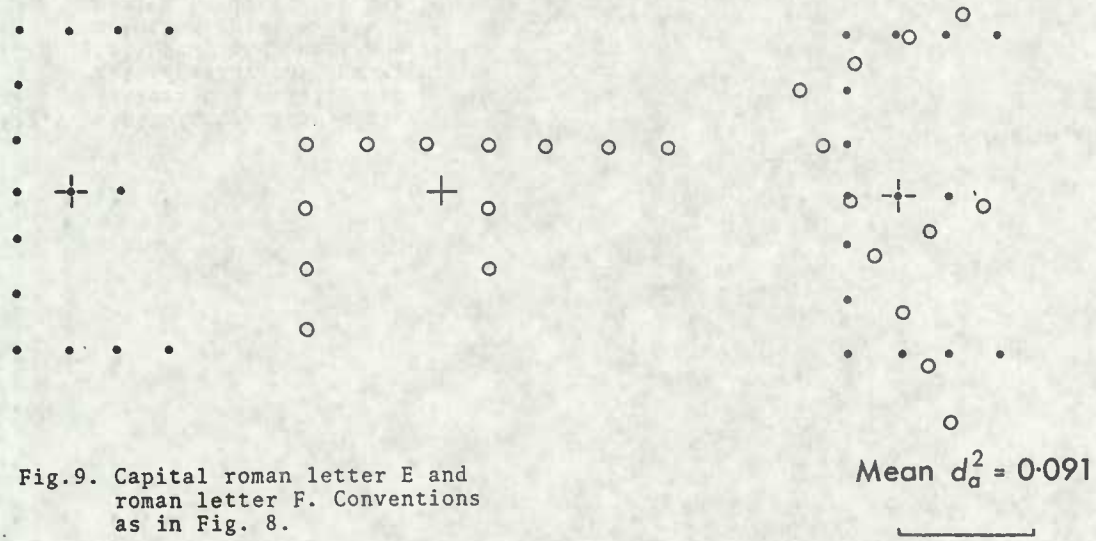

Fig.9. Capital roman letter E and roman letter F. Conventions as in Fig. 8.