CONTINUOUS SYMMETRY FOR SHAPES

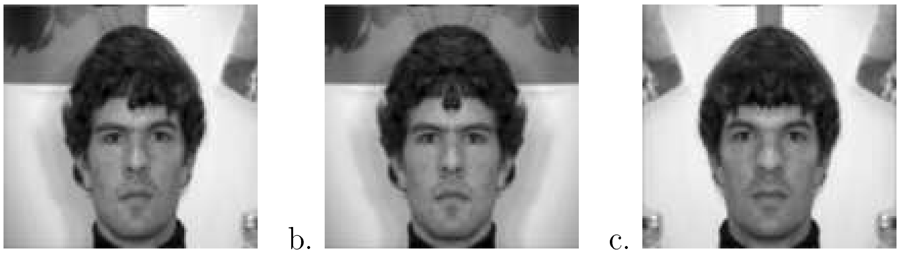
Hagit Zabrodsky and Shmuel Peleg
Institute of Computer Science
and
David Avnir
Departmenet of Organic Chemistry
The Hebrew University of Jerusalem
91904 Jerusalem, Israel

ABSTRACT

Symmetry is treated as a continuous feature, and the Symmetry Distance (SD) of a shape is defined to be the minimum distance required to move points of the original shape in order to obtain a symmetrical shape. This general definition of a symmetry measure enables a comparison of the "amount" of symmetry of different shapes and the "amount" of different symmetries of a single shape. The measure is applicable to any type of symmetry in any dimension. The Symmetry Distance gives rise to a method of reconstructing symmetry of occluded shapes. We extend the method to deal with symmetries of noisy and fuzzy data.

1 Introduction

One of the basic features of shapes and objects is symmetry. Symmetry is considered a pre-attentive feature which enhances recognition and reconstruction of shapes and objects [2]. Symmetry is also an important parameter in physical and chemical processes and is an important criterion in medical diagnosis. However, the exact



a. b. c.

Figure 1: Faces are not perfectly symmetrical.
a) Original image.
b) Right half of original image and its reflection.
c) Left half of original image and its reflection.

mathematical definition of symmetry [4, 6] is inadequate to describe and quantify the symmetries found in the natural world nor those found in the visual world (a classic example is that of faces - see Fig. 1). Furthermore, even perfectly symmetric objects loose their exact symmetry when projected onto the image plane or the retina due to occlusion, perspective transformations, digitization, etc. Thus, although symmetry is usually considered a binary feature, (i.e. an object is either symmetric or it is not symmetric), we view symmetry as a continuous feature where intermediate values of symmetry denote some intermediate "amount" of symmetry. We introduce a "Symmetry Distance" that can measure and quantify all types of symmetries of objects. This measure will enable a comparison of the "amount" of symmetry of different shapes and the "amount" of different symmetries of a single shape.

In Sect. 2 we define the Symmetry Distance and in Sect. 3 we describe a method for evaluting this measure. We then describe features of the symmetry distance including its use in dealing with occluded objects and with noisy data.

2 A Continuous Symmetry Measure - Definition

We define the **Symmetry Distance** (SD) as a the minimum effort required to turn a given shape into a symmetric shape. This is measured by the mean of the square distances each point is moved from its location in the original shape to its location in the symmetric shape. No a priori symmetric reference shape is assumed.

Denote by Ω the space of all shapes of a given dimension, where each shape P is represented by a sequence of n points $\{P_i\}_{i=0}^{n-1}$. We define a metric d on this space as follows:

$$d : \Omega \times \Omega \rightarrow R$$
$$d(P,Q) = d(\{P_i\},\{Q_i\}) = \frac{1}{n}\sum_{i=0}^{n-1} ||P_i - Q_i||^2$$

This metric defines a distance function between every two shapes in Ω.

We define the **Symmetry Transform** ST of a shape P, as the symmetric shape closest to P in terms of the metric d.

The **Symmetry Distance (SD)** of a shape P is now defined as the distance between P and it's Symmetry Transform:

$$SD = d(P, ST(P))$$

The SD of a shape $P = \{P_i\}_{i=0}^{n-1}$ is evaluated by finding the symmetry transform $\hat{P}$ of P (Fig.2) and computing: $\boxed{\text{SD} = \frac{1}{n}\Sigma_{i=0}^{n-1}||P_i - \hat{P}_i||^2}$.

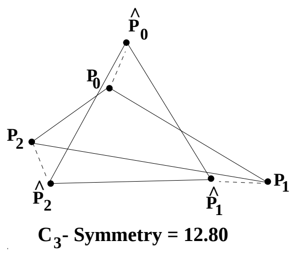


Figure 2: The symmetry transform of $\{P_0, P_1, P_2\}$ is $\{\hat{P}_0, \hat{P}_1, \hat{P}_2\}$. SD $= \frac{1}{3}\Sigma_{i=0}^{2}||P_i - \hat{P}_i||^2$

This definition of the Symmetry Distance implicitly implies invariance to rotation and translation. Normalization of the original shape prior to the transformation additionally allows invariance to scale (Fig. 3). We normalize by scaling the shape so that the maximum distance between points on the contour and the centroid is a given constant (in this paper all examples are given following normalization to 100). The normalization presents an upper bound of on the mean squared distance moved by points of the shape. Thus the SD value is limited in range, where SD=0 for perfectly symmetric shapes (see Appendix in [9]). The general definition of the Sym-

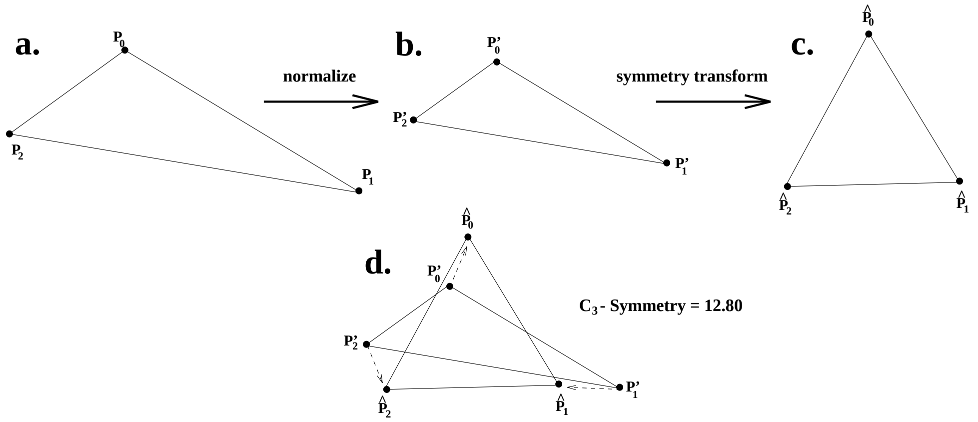


Figure 3: Calculating the Symmetry Distance of a shape:
a) Original shape $\{P_0, P_1, P_2\}$.
b) Normalized shape $\{P'_0, P'_1, P'_2\}$, such that maximum distance to the center of mass is one.
c) Applying the symmetry transform to obtain a symmetric shape $\{\hat{P}_0, \hat{P}_1, \hat{P}_2\}$.
d) SD $= \frac{1}{3}(\|P'_0 - \hat{P}_0\|^2 + \|P'_1 - \hat{P}_1\|^2 + \|P'_2 - \hat{P}_2\|^2)$
SD values are multiplied by 100 for convenience of handling.

metry Distance enables evaluation of a given shape for different types of symmetries (mirror-symmetries, rotational symmetries etc). Moreover, this generalization allows comparisons between the different symmetry types, and allows expressions such as "a shape is more mirror-symmetric than rotationally-symmetric of order two". An additional feature of the Symmetry Distance is that we obtain the symmetric shape which is 'closest' to the given one, enabling visual evaluation of the SD.
An example of a 2D polygon and it's symmetry transforms and SD values are shown in Fig. 4. Note that shape 4e is the most similar to the original shape 4a and, indeed, its SD value is the smallest. In the next Section we describe a geometric algorithm for deriving the Symmetry Transform of a shape. In Sect. 4 we deal with the initial step of representing a shape by a collection of points.

3 Evaluating the Symmetry Transform

In this Section we describe a geometric algorithm for deriving the Symmetry Transform of a shape represented by a sequence of points $\{P_i\}_{i=0}^{n-1}$. In practice we find the Symmetry Transform of the shape with respect to a given point-symmetry

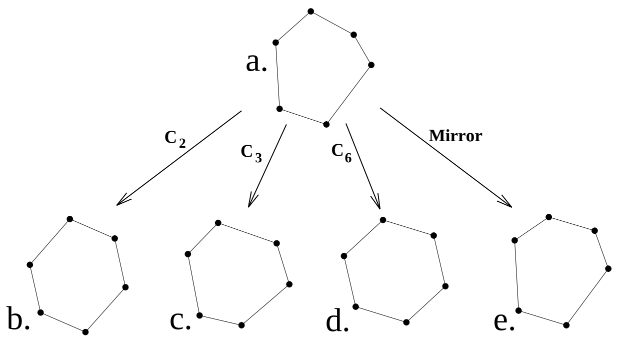


Figure 4: Symmetry Transforms and Symmetry Distances of a 2D polygon.
a) The 2D polygon.
b) Symmetry Transform of (a) with respect to C_2-symmetry (SD = 1.87).
c) Symmetry Transform of (a) with respect to C_3-symmetry (SD = 1.64).
d) Symmetry Transform of (a) with respect to C_6-symmetry (SD = 2.53).
e) Symmetry Transform of (a) with respect to Mirror-symmetry (SD = 0.66).

group. For simplicity and clarity of explanation, we describe the method by using some examples.

Following is a geometrical algorithm for deriving the symmetry transform of a shape P having n points with respect to rotational symmetry of order n (C_n-symmetry). This method transforms P into a regular n-gon, keeping the centroid in place.

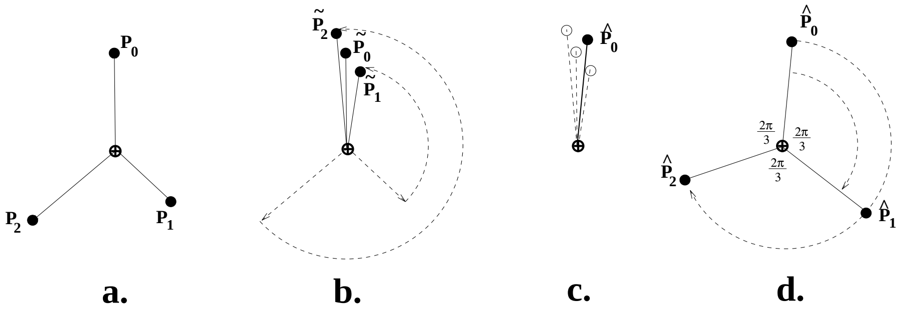


Figure 5: The C_3-symmetry Transform of 3 points.
a) original 3 points $\{P_i\}_{i=0}^{2}$.
b) Fold $\{P_i\}_{i=0}^{2}$ into $\{\tilde{P}_i\}_{i=0}^{2}$.
c) Average $\{\tilde{P}_i\}_{i=0}^{2}$ obtaining $\hat{P}_0 = \frac{1}{3}\sum_{i=0}^{2} \tilde{P}_i$.
d) Unfold the average point obtaining $\{\hat{P}_i\}_{i=0}^{2}$.
The centroid ω is marked by $\oplus$.

Algorithm for finding the C_n-symmetry transform:

1. *Fold* the points $\{P_i\}_{i=0}^{n-1}$ by rotating each point P_i counterclockwise about the centroid by $2\pi i/n$ radians (Fig. 5b).

2. Let $\hat{P}_0$ be the average of the points $\{\tilde{P}_i\}_{i=0}^{n-1}$ (Fig. 5c).

3. *Unfold* the points, obtaining the C_n-symmetric points $\{\hat{P}_i\}_{i=0}^{n-1}$ by duplicating $\hat{P}_0$ and rotating clockwise about the centroid by $2\pi i/n$ radians (Fig. 5d).

The set of points $\{\hat{P}_i\}_{i=0}^{n-1}$ is the symmetry transform of the points $\{P_i\}_{i=0}^{n-1}$. i.e. they are the C_n-symmetric configuration of points closest to $\{P_i\}_{i=0}^{n-1}$ in terms of the metric d defined in Sect. 2.

The common case, however, is that shapes have more points than the order of the symmetry. For symmetry of order n, the folding method can be extended to shapes having a number of points which is a multiple of n. A 2D shape P having qn points is represented as q sets $\{S_r\}_{r=0}^{q-1}$ of n interlaced points $S_r = \{P_{rn+i}\}_{i=0}^{n-1}$. The C_n-symmetry transform of P (Fig. 6) is obtained by applying the above algorithm to each set of n points seperately, where the folding is performed about the centroid of all the points.

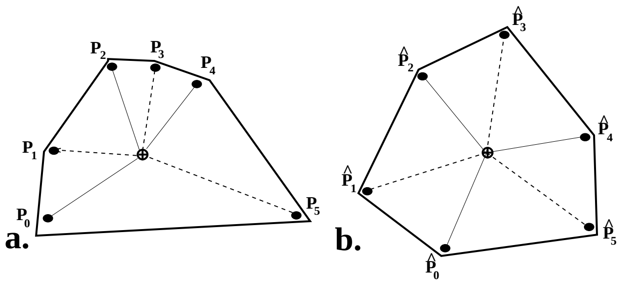


Figure 6: Geometric description of the C_3-symmetry transform for 6 points.
The centroid ω of the points is marked by $\oplus$.
a) The original points shown as 2 sets of 3 points: $S_0 = \{P_0, P_2, P_4\}$ and $S_1 = \{P_1, P_3, P_5\}$.
b) The obtained C_3-symmetric configuration.

The procedure for evaluating the symmetry transform for mirror-symmetry is similar: Given a shape having $m = 2q$ points we divide the points into q pairs of points and given an initial guess of the symmetry axis, we apply the folding/unfolding method as follows (see Fig. 7):

Algorithm for finding the mirror-symmetry transform:

1. for every pair of points $\{P_0, P_1\}$:
   (a) fold - by reflecting across the mirror symmetry axis obtaining $\{\tilde{P}_0, \tilde{P}_1\}$.
   (b) average - obtaining a single averaged point $\hat{P}_0$.
   (c) unfold - by reflecting back across the mirror symmetry axis obtaining $\{\hat{P}_0, \hat{P}_1\}$.
2. minimize over all possible axis of mirror-symmetry.

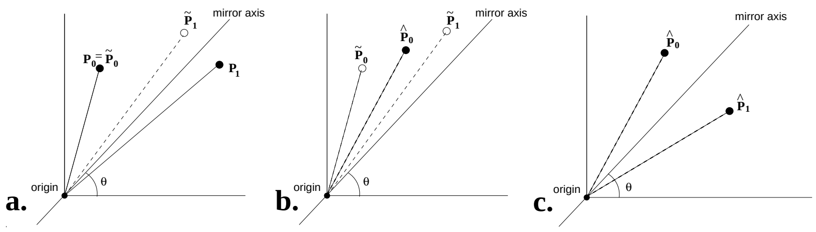


Figure 7: The mirror-symmetry transform of a single pair of points for angle θ, where the centroid of the shape is assumed to be at the origin.
a) The two points $\{P_0, P_1\}$ are folded to obtain $\{\tilde{P}_0, \tilde{P}_1\}$.
b) Points $\tilde{P}_0$ and $\tilde{P}_1$ are averaged to obtain $\hat{P}_0$.
c) $\hat{P}_1$ is obtained by reflecting $\hat{P}_0$ about the symmetry axis.

The minimization performed in step 2 is, in practice, replaced by an analytic solution.

This method extends to **any** finite point-symmetry group G in **any** dimension, where the folding and unfolding are performed by applying the group elements about the centroid.

Given a symmetry group G (having n elements) and given a shape P represented by $m = qn$ points, the symmetry transform of the shape with respect to G-symmetry is obtained as follows:

Algorithm for finding the G-symmetry transform:

- The points are divided into q sets of n points.
- For every set of n points:
  - The points are folded by applying the elements of the G-symmetry group.
  - The folded points are averaged, obtaining a single averaged point.
  - The averaged point is unfolded by applying the inverse of the elements of the G-symmetry group. A G-symmetric set of n points is obtained.
- The above procedure is performed over all possible orientations of the symmetry axis and planes of G. Select that orientation which minimizes the Symmetry Distance value. As previously noted, this minimization is analytic in 2D but requires an iterative minimization process in 3D (except for the 3D mirror-symmetry group where a closed form solution has been derived).

4 Point Selection for Shape Representation

As symmetry has been defined on a sequence of points, representing a given shape by points must precede the application of the symmetry transform. Selection of points influences the value of SD and depends on the type of object to be measured. If a shape is inherently created from points (such as a graph structure or cyclically connected points creating a polygon) we can represent a shape by these points (Fig. 8). This is the case when analysing symmetry of molecules ([10, 9, 7]).

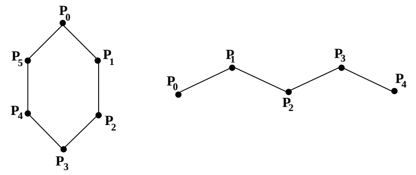


Figure 8: When measuring symmetry of shapes inherently created from points we represent the shape by these points.

There are several ways to select a sequence of points to represent continuous 2D shapes. One such method is selection at **equal distances-** the points are selected along the shape's contour such that the curve length between every pair of adjacent points is equal (Fig. 9).

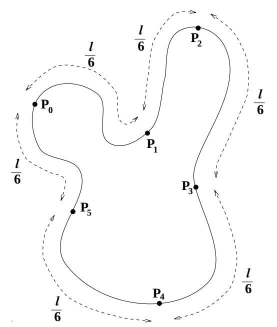


Figure 9: Point selection by equal distance: points are selected along the contour such that each point is equidistant to the next in terms of curve length. In this example six points are distributed along the contour spaced by $\frac{1}{6}$ of the full contour length.

In many cases, however, contour length is not a meaningful measure, as in noisy or occluded shapes. In such cases we propose to select points on a smoothed version of the contour and then project them back onto the original contour. The smoothing of the continuous contour is performed by moving each point on the continuous contour to the centroid of its contour neighborhood. The greater the size of the neighborhood, the greater is the smoothing (see Fig. 10). The level of smoothing can vary and for a high level of smoothing, the resulting shape becomes almost a circle about the centroid ([5]).

We use the following procedure for selection by smoothing:
The original contour is first sampled at very high density and at equal distances along the contour obtaining the sampled points $\{M_j\}$ (Fig. 11a). Following, each sampled point M_j is replaced by the centroid M'_j of a finite number of its neighboring sampled points. The "smoothed" shape is obtained by connecting the centroid points $\{M'_j\}$ (Fig. 11b). A second sampling of points is performed on the smoothed contour, where the points $\{P'_i\}$ are, again, selected at equal distances along the contour. The backprojection is performed by concidering each sampled point P'_i as an interpolated point between two points of the smoothed shape M'_{j_1} and M'_{j_2} (Fig. 11c). The sample point P'_i is backprojected onto the original contour at the point P_i which is the interpolated point between the corresponding sampled points M_{j_1} and M_{j_2} on the original contour

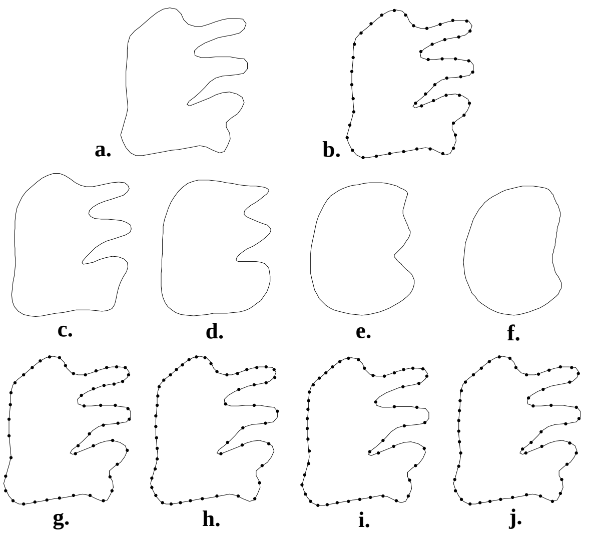


Figure 10: Selection by smoothing
a) Original continuous contour.
b) Points are selected at equal distances along the continuous contour.
c-f) The smoothed shape is obtained by averaging neighboring points of (b). The amount of smoothing depends on the size of the neighborhood. The smoothed shapes (c-f) are obtained when neighborhood includes 5,10,15 and 20 percent of the points, respectively.
g-j) The sampling of points on the original shape using the smoothed shapes (c-f) respectively. Notice that fewer points are selected on the "noisy" part of the contour.

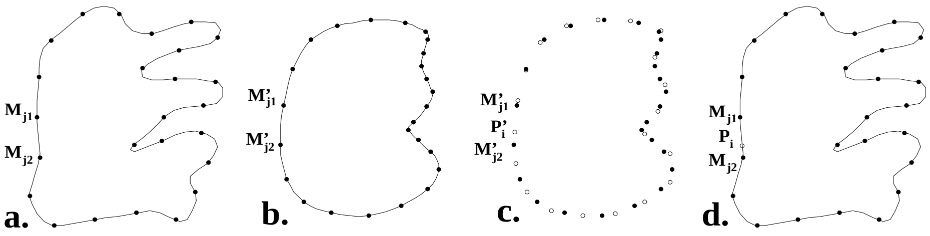


Figure 11: Selection by smoothing (in practice)
a) Continuous contour sampled at equal distances. Points $\{M_j\}$ are obtained.
b) Each sampled point M_j is replaced by the centroid M_j' of a finite number of its neighboring sampled points. The "smoothed" shape is obtained by connecting the centroid points $\{M_j'\}$
c) A second sampling of points is performed at equal distances on the smoothed contour obtaining points $\{P_i'\}$ (white points in Figure). The sample point P_i' is concidered as an interpolation between the two points M_{j_1}' and M_{j_2}'.
d) The sample point P_i' is backprojected onto the original contour at the point P_i which is the interpolated point between the corresponding sampled points M_{j_1} and M_{j_2} on the original contour.

(Fig. 11d). i.e. $P_i = (P_i' - M_{j_1})\frac{(M_{j_2}-M_{j_1})}{(M_{j_2}'-M_{j_1}')} + M_{j_1}$.
The ultimate smoothing is when the shape is smoothed into a circle. In this case,

equal distances on the circular contour is equivalent to equal angles about the center. For maximum smoothing we, therefore, use selection at **equal angles** (Fig. 12) where points are selected on the original contour at equal angular intervals around the centroid.

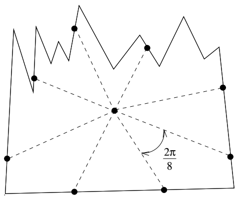


Figure 12: Selection at equal angles: points are distributed along the contour at regular angular intervals around the centroid.

5 Symmetry of Occluded Shapes - Center of Symmetry

When a symmetric object is partially occluded, we use the symmetry distance to evaluate the symmetry of the occluded shapes, locate the center of symmetry and reconstruct the symmetric shape most similar to the unoccluded original.

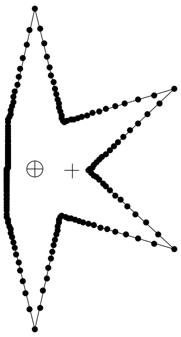


Figure 13: An occluded shape with sampled points selected at equal angles about the center of symmetry (marked by $\oplus$). The symmetry distance obtained using these points is greater than the symmetry distance obtained using points selected at equal angles about the centroid (marked by $+$).

As described in Sect. 4, a shape can be represented by points selected at regular angular intervals about the centroid. Angular selection of points about a point other than the centroid will give a different symmetry distance value. We define the **center of symmetry** of a shape as that point about which selection at equal angles gives the minimum symmetry distance value. When a symmetric shape is complete the center of symmetry coincides with the centroid of the shape. However, the center of symmetry of truncated or occluded objects does not align with its centroid but is closer to the centroid of the complete shape. Thus the center of symmetry of a shape is robust under truncation and occlusion.

To locate the center of symmetry, we use an iterative procedure of gradient descent that converges from the centroid of an occluded shape to the center of symmetry. Denote by *center of selection* the point about which points are selected using selection at equal angles. We initialize the iterative process by setting the centroid as the center of selection. At each step we compare the symmetry value obtained from points selected at equal angles about the center of selection with the symmetry value obtained by selection about points in the center of selection's immediate neighborhood. That point about which selection at equal angles gives minimum symmetry value is set to

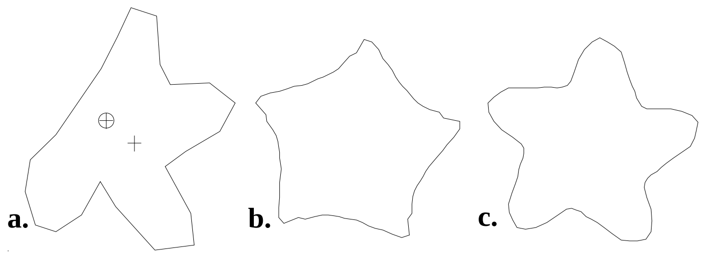


Figure 14: Reconstruction of an occluded shape.
a) Original occluded shape, its centroid (+) and its center of symmetry($\oplus$).
b,c) The closest C_5-symmetric shape following angular selection about the centroid (b) and about the center of symmetry (c). Selection about the centroid gives a featureless shape, while selection about the center of symmetry yields a more meaningful shape.

be the new center of selection. If center of selection does not change, the neighborhood size is decreased. The process is terminated when neighborhood size reaches a predefined minimum size. The center of selection at the end of the process is taken as the center of symmetry.

In the case of occlusions (Fig. 14), the closest symmetric shape obtained by angular selection about the center of symmetry is visually more similar to the original than that obtained by angular selection about the centroid. We can reconstruct the symmetric shape closest to the unoccluded shape by obtaining the symmetry transform of the occluded shape using angular selection about the center of symmetry (see Fig. 14c). In Fig. 15 the center of symmetry and the closest symmetric shapes were found for several occluded flowers.

The process of reconstructing the occluded shape can be improved by altering the method of evaluating the symmetry of a set of points. As described in Sect. 3, the symmetry of a set of points is evaluated by folding, averaging and unfolding about the centroid of the points. We alter the method as follows:

1. The folding and unfolding (steps 1 and 3) will be performed about the center of selection rather than about the centroid of the points.

2. Rather than averaging the folded points (step 2), we apply other robust clustering methods. In practice we average over the folded points, drop the points farthest from the average and then reaverage (see Fig. 16).

The improvement in reconstruction of an occluded shape is shown in Fig. 17. This method improves both shape and localization of the reconstruction. Assuming that the original shape was symmetric, this method can reconstruct an occluded shape very accurately.

6 Symmetry of Points with Uncertain Locations

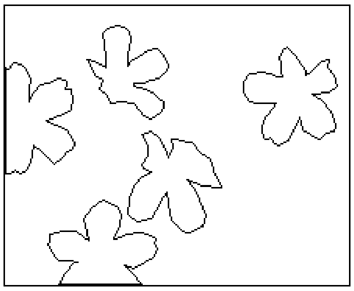
In most cases, sensing processes do not have absolute accuracy and the location of each point in a sensed pattern can be given only as a probability distribution. Given sensed points with such uncertain locations, the following properties are of interest:

- The most probable symmetric configuration represented by the sensed points.
- The probability distribution of symmetry distance values for the sensed points.

6.1 *The Most Probable Symmetric Shape*

Fig. 18a shows a configuration of points whose locations are given by a normal distribution function. The dot represents the expected location of the point and the rectangle represents the standard deviation marked as rectangles having width and length proportional to the standard deviation. In this section we briefly describe a method of evaluating the most probable symmetric shape under the Maximum Likelihood criterion ([3]) given the sensed points. Detailed derivations and proofs are given in Appendix A.1. For simplicity we describe the method with respect to rotational symmetry of order n (C_n-symmetry). The solution for mirror symmetry or any other symmetry is similar (see Appendix A.2).

Given n ordered points in 2D whose locations are given as normal probability distributions with expected location P_i and covariance matrix Λ_i:
$Q_i \sim \mathcal{N}(P_i, \Lambda_i) \quad i = 0 \ldots n-1$, we find the C_n-symmetric configuration of points at locations $\{\hat{P}_i\}_0^{n-1}$ which is optimal under the Maximum Likelihood criterion ([3]).
Denote by ω the centeroid of the most probable C_n-symmetric set of locations $\hat{P}_i$:

a.  b. 

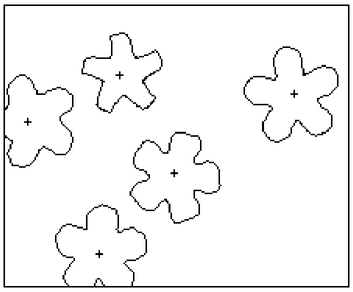
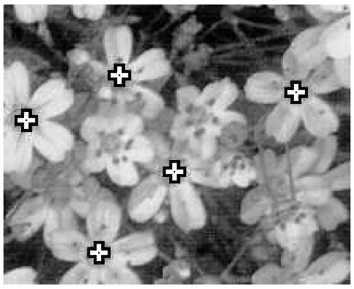
c.  d. 

Figure 15: Application example.
a) A collection of occluded asymmetric flowers.
b) Contours of the occluded flowers were extracted manually.
c) The closest symmetric shapes and their center of symmetry.
d) The center of symmetry of the occluded flowers are marked by '+'.

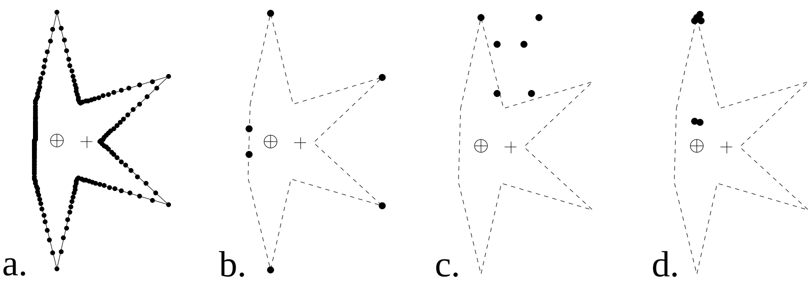


Figure 16: Improving the averaging of folded points
a) An occluded shape with points selected using angular selection about the center of symmetry (marked as $\oplus$).
b) A single set (orbit) of the selected points of a) is shown.
c) folding the points about the centroid (marked as $+$), points are clustered sparsely.
d) folding the points about the center of symmetry, points are clustered tightly. Eliminating the extremes (two farthest points) and averaging results in a smaller averaging error and better reconstruction.

$\omega = \frac{1}{n}\sum_{i=0}^{n-1}\hat{P}_i$. The point ω is dependent on the location of the measurements (P_i) and on the probability distribution associated with them (Λ_i). Intuitively, ω is positioned at that point about which the folding (described below) gives the tightest cluster of points with small uncertainty (small s.t.d.). We assume for the moment that ω is given. A method for finding ω is derived in Appendix A.1. We use a variant of the folding method which was described in Sect. 3 for evaluating C_n-symmetry of a set of points:

1. The n measurements $Q_i \sim \mathcal{N}(P_i, \Lambda_i)$ are *folded* by rotating each measurement Q_i by $2\pi i/n$ radians about the point ω. A new set of measurements $\tilde{Q}_i \sim \mathcal{N}(\tilde{P}_i, \tilde{\Lambda}_i)$ is obtained (see Fig. 18b).

2. The folded measurements are *averaged* using a weighted average, obtaining a single point $\hat{P}_0$. Averaging is done by considering the n folded measurements $\tilde{Q}_i$ as n measurements of a single point and $\hat{P}_0$ represents the most probable

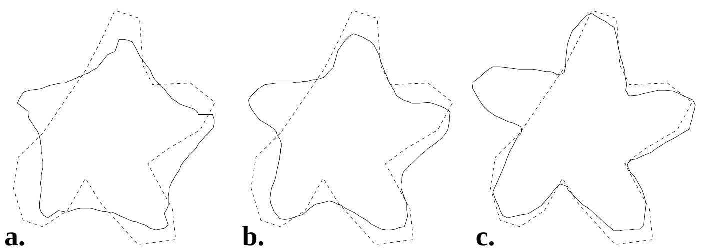


Figure 17: Reconstruction of an occluded almost symmetric shape.
The original shape is shown as a dashed line and the reconstructed shape as a solid line.
a) The closest symmetric shape following angular selection about the centroid.
b) The closest symmetric shape following angular selection about the center of symmetry.
c) The closest symmetric shape following angular selection about the center of symmetry and altered symmetry evaluation (see text).

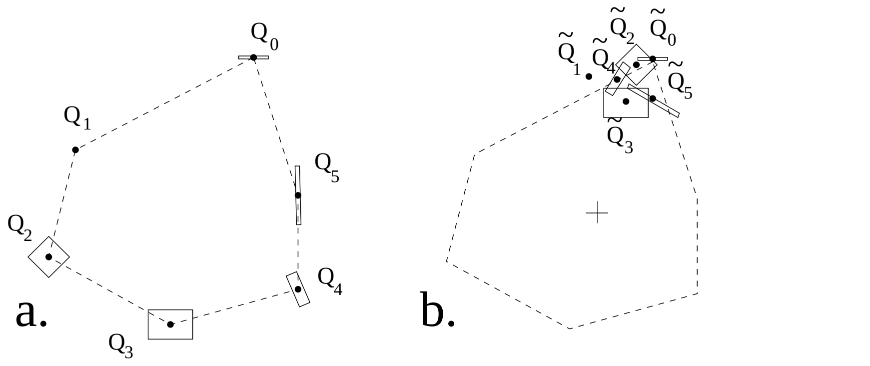


Figure 18: Folding measured points.
a) A configuration of 6 measuremented points $Q_0 \ldots Q_5$. The dot represents the expected location of the point and the rectangle represents the standard deviation marked as rectangles having width and length proportional to the standard deviation.
b) Each measurement Q_i was rotated by $2\pi i/6$ radians about the centroid of the expected point locations (marked as '+') obtaining measurement $\tilde{Q}_0 \ldots \tilde{Q}_5$.

location of that point under the Maximum Likelihood criterion.

$$\hat{P}_0 - \omega = (\sum_{j=0}^{n-1} \tilde{\Lambda}_j^{-1})^{-1} \sum_{i=0}^{n-1} \tilde{\Lambda}_i^{-1} \tilde{P}_i - \omega$$

3. The "average" point $\hat{P}_0$ is *unfolded* as described in Sect. 3 obtaining points $\{\hat{P}_i\}_{i=0}^{n-1}$ which are perfectly C_n-symmetric.

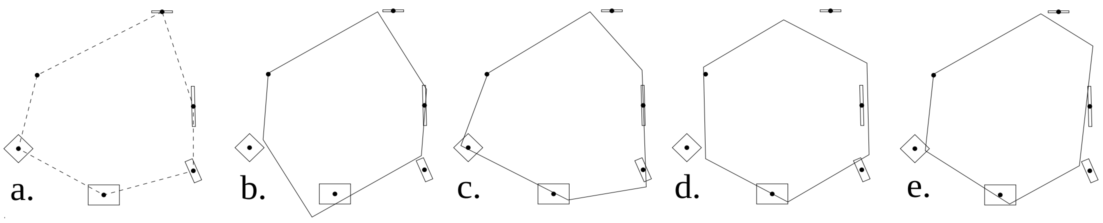


Figure 19: The most probable symmetric shape.
a) A configuration of 6 measured points.
b) The most probable symmetric shapes with respect to C_2-symmetry.
c) The most probable symmetric shapes with respect to C_3-symmetry.
d) The most probable symmetric shapes with respect to C_6-symmetry.
e) The most probable symmetric shapes with respect to mirror-symmetry.

When we are given $m = qn$ measurements, we find the most probable C_n-symmetric configuration of points, similar to the folding method of Sect. 3. The m measurements $\{Q_i\}_{i=0}^{m-1}$, are divided into q interlaced sets of n points each, and the folding method as described above is applied seperately to each set of measurements. Derivations and proof of this case is also found in Appendix A.1.
Several examples are shown in Figure 19, where for a given set of measurements, the

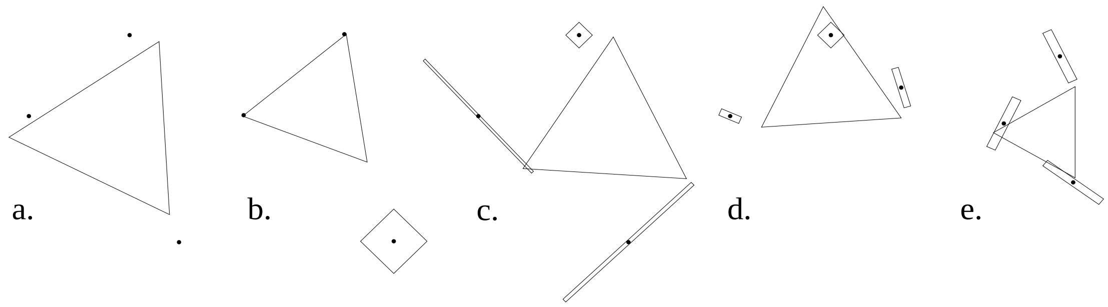


Figure 20: The most probable C_3-symmetric shape for a set of measurements after varying the probability distribution and expected locations of the measurements.
a-c) Changing the uncertainty (s.t.d.) of the measurements.
d-e) Changing both the uncertainty and the expected location of the measurements.

most probable symmetric shapes were found. Fig. 20 shows an example of varying the probability distribution of the measurements on the resulting symmetric shape.

6.2 *The Probability Distribution of Symmetry Values*

Fig. 21a displays a Laue photograph ([1]) which is an interefrence pattern created by projecting X-ray beams onto crystals. Crystal quality is determined by evaluating the symmetry of the pattern. In this case the interesting feature is not the closest symmetric configuration, but the probability distribution of the symmetry distance values.

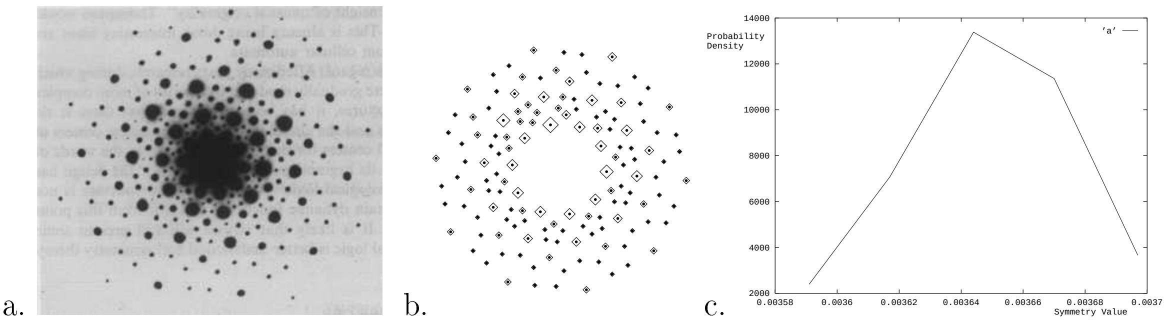


Figure 21: Probability distribution of symmetry values
a) Interference pattern of crystals.
b) Probability distribution of point locations corresponding to a.
c) Probability distribution of symmetry distance values with respect to C_{10}-symmetry was evaluated as described in Appendix A.3. Expectation value = 0.003663.

In Fig. 22 we display distributions of the symmetry distance value for various measurements. As expected, the distribution of symmetry distance values becomes broader as the uncertainties (the variance of the distribution) of the measurements increase.

Consider the configuration of 2D measurements given in Fig. 18a. Each measurement Q_i is a normal probability distribution $Q_i \sim \mathcal{N}(P_i, \Lambda_i)$. We assume the centroid of the expectation of the measurements is at the origin. The probability distribution of the symmetry distance values of the original measurements is equivalent to the probability distribution of the location of the "average" point ($\hat{P}_0$) given the folded measurements as obtained in Step 1 and Step 2 of the algorithm in Sect 6.1. It is shown in Appendix A.3 that this probability distribution is a χ^2 distribution of order $n-1$. However, we can approximate the distribution by a gaussian distribution. Details of the derivation are given in Appendix A.3.

In Fig. 21 we display distributions of the symmetry distance value as obtained for the Laue photograph given in Fig. 21a. In this example we considered every dark patch as a measured point with variance proportional to the size of the patch. Thus in Fig. 21b the rectangles which are proportional in size to the corresponding dark patches of Fig. 21a, represent the standard deviation of the locations of point measurements. Note that a different analysis could be used in which the variance of the measurement location is taken as inversely proportional to the size of the dark patch.

7 Conclusion

We view symmetry as a continuous feature and define a Symmetry Distance (SD) of shapes. The general definition of the Symmetry Distance enables a comparison of

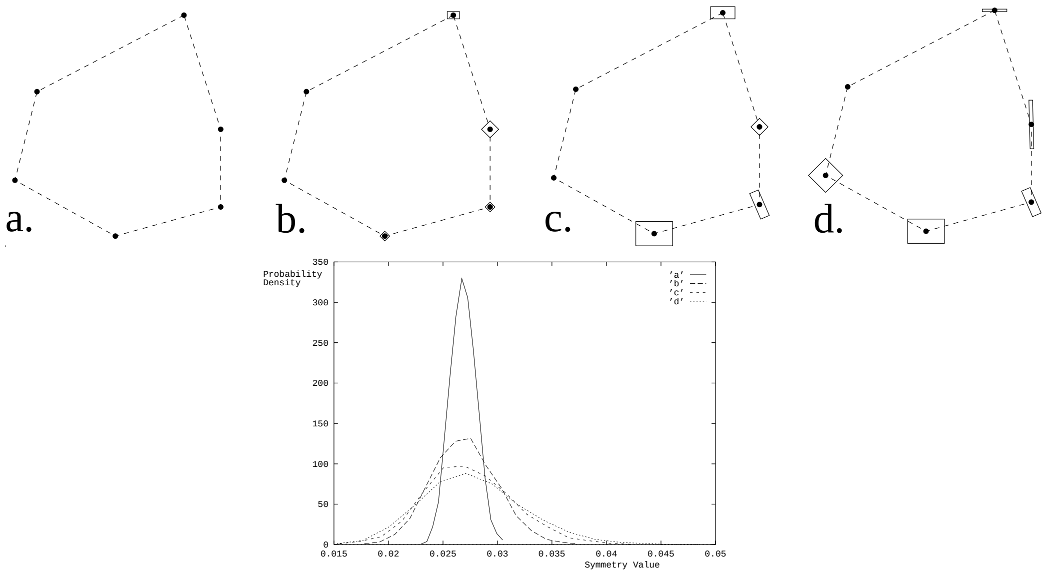


Figure 22: Probability distribution of the symmetry distance value as a function of the variance of the measured points.
a-d) Some examples of configurations of measured points.
e) Probability distribution of symmetry distance values with respect to C_6-symmetry for the configurations a-d.

the "amount" of symmetry of different shapes and the "amount" of different symmetries of a single shape. Furthermore, the SD is associated with the symmetric shape which is 'closest' to the given one, enabling visual evaluation of the SD. Several applications were described including reconstruction of occluded shapes, finding face orientation and finding locally symmetric regions in images. We also described how we deal with uncertain data, i.e. with a configuration of measurements representing the probability distribution of point location. The methods described here can be easily extended to higher dimensions and to more complex symmetry groups. Further extensions will deal with other symmetry classes such as planar symmetry (including translatory symmetry and fractals). Additional work has been done on evaluating reflective symmetry (and chirality) of graph structures ([7, 8]). Current work is expanding the method described here to deal with skewed and projected symmetries.

Appendix

A Uncertain Point Locations

A.1 *The most Probable C_n-Symmetric Shape*

In Sect 6.1 we described a method of evaluating the most probable symmetric shape given a set of measurements. In this Section we derive mathematically and prove the method. For simplicity we derive the method with respect to rotational symmetry of order n (C_n-symmetry). The solution for mirror symmetry is similar (see Appendix A.2).

Given n points in 2D whose positions are given as normal probability distributions: $Q_i \sim \mathcal{N}(P_i, \Lambda_i), \quad i = 0 \ldots n-1$, we find the C_n-symmetric configuration of points $\{\hat{P}_i\}_0^{n-1}$ which is most optimal under the Maximum Likelihood criterion ([3]).

Denote by ω the center of mass of $\hat{P}_i$: $\quad \omega = \frac{1}{n}\sum_{i=0}^{n-1} \hat{P}_i$.

Having that $\{\hat{P}_i\}_0^{n-1}$ are C_n-symmetric, the following is satisfied:

$$\hat{P}_i = R_i(\hat{P}_0 - \omega) + \omega \tag{1}$$

for $i = 0 \ldots \quad n-1$ where R_i is a matrix representing a rotation of $2\pi i/n$ radians.

Thus, given the measurements $Q_0, \ldots, Q_{n-1}$ we find the most probable $\hat{P}_0$ and ω. We find $\hat{P}_0$ and ω that maximize $\mathrm{Prob}(\{\mathrm{P_i}\}_{\mathrm{i=0}}^{\mathrm{n-1}} \quad | \quad \omega, \hat{\mathrm{P}}_0)$ under the symmetry constraints of Equation 1.

Considering the normal distribution we have:

$$\prod_{i=0}^{n-1} k_i \exp(-\frac{1}{2}(\hat{P}_i - P_i)^t \Lambda_i^{-1}(\hat{P}_i - P_i)$$

where $k_i = \frac{1}{2}\pi \mid \Lambda_i \mid^{1/2}$. Having log being a monotonic function, we maximize:

$$\log \prod_{i=0}^{n-1} k_i \exp(-\frac{1}{2}(\hat{P}_i - P_i)^t \Lambda_i^{-1}(\hat{P}_i - P_i)$$

Thus we find those parameters which maximize:

$$-\frac{1}{2}\sum_{i=0}^{n-1}(\hat{P}_i - P_i)^t \Lambda_i^{-1}(\hat{P}_i - P_i)$$

under the symmetry constraint of Equation 1.

Substituting Equation 1, taking the derivative with respect to $\hat{P}_0$ and equating to zero we obtain:

$$\underbrace{(\sum_{i=0}^{n-1} R_i^t \Lambda_i^{-1} R_i)\,\hat{P}_0}_{A} + \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_i^{-1}(I - R_i)\,\omega}_{B} = \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_i^{-1} P_i}_{E} \tag{2}$$

Note that $R_0 = I$ where I is the identity matrix.
When the derivative with respect to ω is zero:

$$\underbrace{(\sum_{i=0}^{n-1}(I - R_i)^t \Lambda_i^{-1} R_i)\,\hat{P}_0}_{C} + \underbrace{\sum_{i=0}^{n-1}(I - R_i)^t \Lambda_i^{-1}(I - R_i)\,\omega}_{D} = \underbrace{\sum_{i=0}^{n-1}(I - R_i)^t \Lambda_i^{-1} P_i}_{F} \tag{3}$$

From Equation 2 we obtain

$$\hat{P}_0 - \omega = (\sum_{j=0}^{n-1} R_j^t \Lambda_j^{-1} R_j)^{-1} \sum_{i=0}^{n-1}(R_i^t \Lambda_i^{-1} R_i) R_i^t P_i$$

Which gives the folding method described in Sect 6.1, where $R_i^t P_i$ is the location of the folded measurement (denoted $\tilde{P}_i$ in the text) and $R_i^t \Lambda_i^{-1} R_i$ is its probability distribution (denoted $\tilde{\Lambda}_i$ in the text). The factor $(\sum_{j=0}^{n-1} R_j^t \Lambda_j^{-1} R_j)$ is the normalization factor.

Reformulating Eqs. 2 and 3 in matrix formation we obtain:

$$\underbrace{\begin{pmatrix} A & B \\ C & D \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} \hat{P}_0 \\ \omega \end{pmatrix}}_{V} = \underbrace{\begin{pmatrix} E \\ F \end{pmatrix}}_{Z}$$

Noting that U is symmetric we solve by inversion $V = U^{-1}Z$ obtaining the parameters ω,$\hat{P}_0$ and obtaining the most probable C_n-symmetric configuration, given the measurements $\{Q_i\}_{i=0}^{n-1}$.

Similar to the representation in Sect. 3, given $m = qn$ measurements $\{Q_i\}_{i=0}^{m-1}$, we consider them as q sets of n interlaced measurements: $\{Q_{iq+j}\}_{i=0}^{n-1}$ for $j = 0 \ldots q-1$.

The derivations given above are applied to each set of n measurements separately, inorder to obtain the most probable C_n-symmetric set of points $\{\hat{P}_i\}_{i=0}^{m-1}$. Thus the symmetry constraints that must be satisfied are:

$$\hat{P}_{iq+j} = R_i(\hat{P}_j - \omega) + \omega$$

for $j = 0 \ldots \; q-1$ and $i = 0 \ldots \; n-1$ where, again, R_i is a matrix representing a rotation of $2\pi i/n$ radians and ω is the centroid of all points $\{\hat{P}_i\}_{i=0}^{m-1}$.
As derived in Equation 2, we obtain for $j = 0 \ldots q-1$:

$$\underbrace{(\sum_{i=0}^{n-1} R_i^t \Lambda_{iq+j}^{-1} R_i)}_{A_j} \hat{P}_j + \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_{iq+j}^{-1} (I - R_i)\, \omega}_{B_j} = \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_{iq+j}^{-1} P_{iq+j}}_{E_j} \tag{4}$$

and equating to zero, the derivative with respect to ω, we obtain, similar to Equation 3:

$$\sum_{j=0}^{q-1} \underbrace{(\sum_{i=0}^{n-1} (I-R_i)^t \Lambda_{iq+j}^{-1} R_i)}_{C_j} \hat{P}_j + \underbrace{\sum_{j=0}^{q-1} \sum_{i=0}^{n-1} (I-R_i)^t \Lambda_{iq+j}^{-1} (I-R_i)\, \omega}_{D} = \underbrace{\sum_{j=0}^{q-1} \sum_{i=0}^{n-1} (I-R_i)^t \Lambda_{iq+j}^{-1} P_{iq+j}}_{F} \tag{5}$$

Reformulating Eqs. 4 and 5 in matrix formation we obtain:

$$\underbrace{\begin{pmatrix} A_0 & & & & B_0 \\ & A_1 & & & B_1 \\ & & \ddots & & \vdots \\ & & & A_{q-1} & B_{q-1} \\ C_0 & C_1 & \cdots & C_{q-1} & D \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} \hat{P}_0 \\ \hat{P}_1 \\ \vdots \\ \hat{P}_{q-1} \\ \omega \end{pmatrix}}_{V} = \underbrace{\begin{pmatrix} E_0 \\ E_1 \\ \vdots \\ E_{q-1} \\ F \end{pmatrix}}_{Z}$$

Noting that U is symmetric we solve by inversion $V = U^{-1}Z$ and obtain the parameters ω and $\{\hat{P}_j\}_{j=0}^{q-1}$, and obtain the most probable C_n-symmetric configuration, $\{\hat{P}_j\}_{j=0}^{m-1}$ given the measurements $\{Q_i\}_{i=0}^{m-1}$.

A.2 *The Most Probable Mirror Symmetric Shape*

In Sect 6.1 we described a method for finding the most probable rotationaly symmetric shape given measurements of point location. The solution for mirror symmetry is similar. In this case, given m measurements (where $m = 2q$), the unknown parameters are $\{\hat{P}_j\}_{j=0}^{q-1}, \omega$ and θ where θ is the angle of the reflection axis. However these parameters are redundant and we reduce the dimensionality of the problem by replacing the 2 dimensional ω with the one dimensional x_0 representing the x-coordinate at which the reflection axis intersects the x-axis. Additionally we replace R_i the rotation matrix with $R = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$ the reflection about an axis at an angle θ to the x-axis. The angle θ is found analyticaly (see [10]) thus the dimensionality of the problem is $2q+1$ (rather than $2q+2$) and elimination of the last row and column of matrix U (see Sect 6.1) allows an inverse solution as in the rotational symmetry case.

A.3 *Probability Distribution of Symmetry Values*

In this section we derive mathematically the probability distribution of symmetry distance values obtained from a set of n measurements in 2D: $Q_i \sim \mathcal{N}(P_i, \Lambda_i) \quad i = 0 \ldots n-1$ with respect to C_n-symmetry (see Sect. 6.2).

Denote by X_i the 2-dimensional random variable having a normal distribution equal to that of measurement $\tilde{Q}_i$ i.e.

$$\begin{aligned} E(X_i) &= R_i P_i \\ \mathrm{Cov}(X_i) &= R_i \Lambda_i R_i^t \end{aligned}$$

where R_i denotes (as in Sect. 3) the rotation matrix of $2\pi i/n$ radians.

Denote by Y_i the 2-dimensional random variable:

$$Y_i = X_i - \frac{1}{n}\sum_{j=0}^{n-1} X_j$$

in matrix notation:

$$\underbrace{\begin{pmatrix} Y_0 \\ \vdots \\ Y_{n-1} \end{pmatrix}}_{\mathbf{Y}} = A \underbrace{\begin{pmatrix} X_0 \\ \vdots \\ X_{n-1} \end{pmatrix}}_{\mathbf{X}}$$

or $\mathbf{Y} = A\mathbf{X}$ where $\mathbf{Y}$ and $\mathbf{X}$ are of dimension $2n$ and A is the $2n \times 2n$ matrix:

$$A = \tfrac{1}{n} \begin{pmatrix} n-1 & 0 & -1 & 0 & -1 & \cdots \\ 0 & n-1 & 0 & -1 & 0 & \cdots \\ -1 & 0 & \ddots & 0 & -1 & \cdots \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & \cdots & & n-1 \end{pmatrix}$$

And we have

$$E(\mathbf{X}) = \begin{pmatrix} E(X_0) \\ \vdots \\ E(X_{n-1}) \end{pmatrix} \quad \mathrm{Cov}(\mathbf{X}) = \begin{pmatrix} \mathrm{Cov}(X_0) & & \\ & \ddots & \\ & & \mathrm{Cov}(X_{n-1}) \end{pmatrix}$$

$$E(\mathbf{Y}) = AE(\mathbf{X}) \qquad \mathrm{Cov}(\mathbf{Y}) = A\mathrm{Cov}(\mathbf{X})A^t$$

The matrix $A\mathrm{Cov}(\mathbf{X})A^t$, being symmetric and positive definite, we find the $2n \times 2n$ matrix S diagonalizing $\mathrm{Cov}(\mathbf{Y})$ i.e.

$$SA\mathrm{Cov}(\mathbf{X})A^t S^t = D$$

where D is a diagonal matrix (of rank $2(n-1)$).

Denote by $\mathbf{Z} = (Z_0, \ldots, Z_{n-1})^t$ the $2n$-dimensional random variable $SA\mathbf{X}$.

$$\begin{array}{ll} E(\mathbf{Z}) & = SAE(\mathbf{X}) \\ \mathrm{Cov}(\mathbf{Z}) & = SA\mathrm{Cov}(\mathbf{X})A^tS^t = D \end{array}$$

Thus the random variables Z_i that compose $\mathbf{Z}$ are independent and, being linear combinations of X_i, they are of normal distribution.

The symmetry distance, as defined in Sect. 3, is equivalent, in the current notations, to $s = \mathbf{Y}^t\mathbf{Y}$. Having S orthonormal we have

$$s = (SA\mathbf{X})^t SA\mathbf{X} = \mathbf{Z}^t\mathbf{Z}$$

If $\mathbf{Z}$ were a random variable of standard normal distribution, we would have s being of a χ^2 distribution of order $2(n-1)$. In the general case Z_i are normally distributed but not standard and $\mathbf{Z}$ cannot be standardized globally. We approximate the distribution of s as a normal distribution with

$$\begin{array}{ll} E(s) & = E(\mathbf{Z})^t E(\mathbf{Z}) + \mathrm{trace} D^t D \\ \mathrm{Cov}(s) & = 2\mathrm{trace}(D^t D)(D^t D) + 4E(\mathbf{Z})^t D^t D E(\mathbf{Z}) \end{array}$$

References

[1] J.L. Amoros, M.J. Buerger, and M.Canut de Amoros. *The Laue Method.* Academic Press, New York, 1975.

[2] F. Attneave. Symmetry information and memory for patterns. *American Journal of Psychology*, 68:209–222, 1955.

[3] M. H. DeGroot. *Probability and Statistics.* Addison-Wesley, Reading, MA, 1975.

[4] W. Miller. *Symmetry Groups and their Applications.* Academic Press, London, 1972.

[5] F. Mokhatarian and A. Mackworth. A theory of multiscale, curvature-based shape representation for planar curves. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14:789–805, 1992.

[6] H. Weyl. *Symmetry.* Princeton Univ. Press, 1952.

[7] H. Zabrodsky and D. Avnir. Measuring symmetry in structural chemistry. In I. Hargittai, editor, *Advanced Molecular Structure Research*, volume 1. 1993.

[8] H. Zabrodsky, S. Peleg, and D. Avnir. Continuous symmetry measures, IV: Chirality. In Preparation.

[9] H. Zabrodsky, S. Peleg, and D. Avnir. Continuous symmetry measures II: Symmetry groups and the tetrahedron. *J. Am. Chem. Soc.*, 115:8278–8289, 1993.

[10] H. Zabrodsky, S. Peleg, and D. Avnir. Continuous symmetry measures. *J. Am. Chem. Soc.*, 114:7843–7851, Sept 1992.