American Journal of Botany 79(1): 69–76. 1992.

LANDMARK EIGENSHAPE ANALYSIS: HOMOLOGOUS CONTOURS: LEAF SHAPE IN SYNGONIUM (ARACEAE)[1]

THOMAS S. RAY

School of Life and Health Sciences, University of Delaware, Newark, Delaware 19716

A new morphometric technique suitable for the description of leaf shape is presented that combines features of both outline and landmark approaches. The concept of homologous point landmarks is extended to higher dimensional features: homologous contours. This allows an integrated approach to morphometric analysis, which treats contours together with points while adhering to the practice of comparing only homologous features. The technique modifies eigenshape analysis by defining homologous contours as any contour bounded by two landmarks, containing no landmark on the contour. The contour need not be a closed curve and may occur in the interior as well as on the outline of the specimen. Collections of contours may be treated, and need not be contiguous. Homologous contours are divided into equal numbers of equal-length steps on each specimen, and the angles representing each step as well as the proportional lengths of the contours are the data that enter into the analysis. The technique can be used to discriminate the shapes of leaves or other organs in systematic, ecological, or developmental studies.

Leaf shape can be an important character in a wide variety of botanical studies, ranging from systematic to ecological or developmental. Recent analyses of leaf shape in this journal include those of McLellan (1990) and Jensen (1990). In general, shape has been an underutilized character because it can be difficult to quantify. New methods, combining image-analysis technology, advanced morphometric analysis, and growing computer power are making shape a more accessible character. This paper describes a new morphometric analytical technique as applied to the description of variation of leaf shape in *Syngonium podophyllum.*

Convincing arguments have been made that the meaningful biological basis for the comparison of shapes depends on comparison of homologous structures (Bookstein et al., 1982, 1985). In practice, homology in morphometrics has generally been applied exclusively to point landmarks, and a split exists between techniques that utilize landmark data and techniques that utilize outline data. Outline techniques generally represent an object as a closed curve with at most one landmark, and make a "mathematical homology" between points a proportion p around each outline. However, this supposed homology is usually violated by real data. Consider a comparison of the outlines of two leaves. On one leaf, two adjacent landmarks such as the point of attachment of the petiole and a leaf tip, may be separated by a proportion p of the total length of the outline. However, on a second leaf, these same two landmarks may be separated by a distance of $p + \lambda$. Thus the two landmarks do not line up in the comparison, and the mathematical homology is violated.

The importance of homologous landmarks in morphometric analysis is generally recognized; however, pure landmark techniques do not consider outlines and other contours that are significant aspects of shape. Bookstein (1978) has argued that schemes involving only landmarks are inadequate for the analysis of curving forms. On the other hand, Bookstein et al. (1982, 1985) argued that "half the information of the data base has been intentionally discarded" by outline techniques that ignore landmarks. One could counter that many landmark techniques throw away the other half by ignoring contours.

Outline and landmark techniques need not be mutually exclusive and can be quite easily and simply combined. This paper presents a technique, "landmark eigenshape analysis," that combines landmark and contour data into an integrated analysis providing a very efficient description of shape. This integration is accomplished through a modification of "eigenshape analysis" (Lohmann, 1983; Lohmann and Schweitzer, 1990).

MATERIALS AND METHODS

Data acquisition—Leaf samples were collected from *Syngonium podophyllum* Schott var. *peliocladum* (Schott) Croat (voucher: Barry Hammel 12787, Missouri Botanical Garden) in May 1988 at Finca El Bejuco in the lowland rain forests of the Sarapiquí region of Costa Rica. Leaves were collected in four groups totaling 86 leaves. One group consisted of 29 leaves gathered from individual plants selected to represent the full range of variation in leaf form for the species. Each of the other three groups was collected along shoots of individual plants exhibiting marked heteroblasty.

The dried leaves were digitized using an image acquisition system. The leaves were placed with the underside facing the camera, and the outline was traced counterclockwise, with the point of insertion of the petiole as the starting point. The image acquisition system then recorded the x, y coordinates of 128 roughly equally spaced points around the outline, saving these data in a file with the letter and number label for each leaf. The output file lists coordinates using an arbitrary scale, but it also includes a scaling factor for each sample so that the arbitrary

---

[1] Received for publication 18 December 1990; revision accepted 20 September 1990.

The author thanks Carlos Campos for assistance; Chris Meacham, Pat Lohmann, Jack Fisher, and Tim Dickinson for critical reviews of the manuscript; John Schuenemeyer and Devon Mehrota of the University of Delaware Statistical Laboratory for consultations; Peter Schweitzer and Norman MacLeod for generously providing the source code for their eigenshape analysis programs, which are modifications of original programs written by Lohmann and Denham; and Norman MacLeod for making his image analysis system available. This study was supported in part by NSF grant BSR 8822576. Contribution No. 152 from the Ecology Program, School of Life and Health Sciences, University of Delaware.

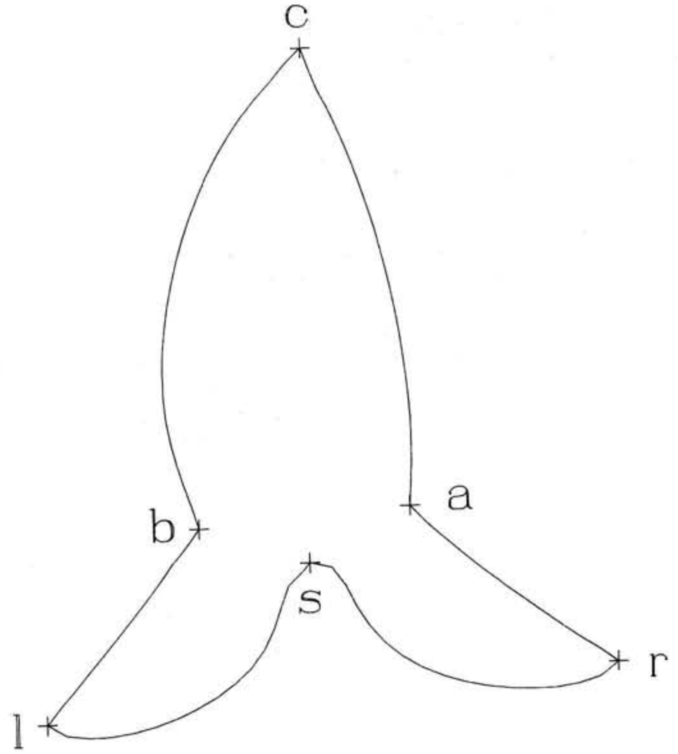

Fig. 1. The average leaf outline, with the four landmarks, s, r, c, and l, and the two pseudolandmarks, a and b, marked and labeled.

units can be adjusted to measurements in the units defined at the time the data were acquired.

Landmark eigenshape analysis—The basic principal of this approach is to identify "homologous contours" bounded by homologous landmarks and to divide these contours into a number of equal length steps, which can be represented by their angles in the plane. This may be viewed as an extension of eigenshape analysis as developed by Lohmann (1983) and Schweitzer, Kaesler, and Lohmann (1986), who applied it in the case of outlines with no landmarks. Landmark eigenshape analysis is in essence the method used by Lohmann and Schweitzer (1990), who traced each specimen with two outlines at right angles in order to provide a "quasi-three-dimensional" representation. Their application uses one homologous contour per outline and two outlines per specimen. In the application presented here, there is only one outline per specimen, but each outline is divided into six homologous contours, demarcated by six landmarks.

Homologous contours are defined in this application as illustrated in Fig. 1, which is the average of the 86 leaf outlines of the data set. In this outline, there are four clearly defined landmarks: the point of insertion of the petiole 's', the tip of the right lobe 'r', the tip of the central lobe 'c', and the tip of the left lobe 'l'. In addition there are two pseudolandmarks: the base of the sinus between the central lobe and the right lobe 'a', and the base of the sinus between the left lobe and the central lobe 'b'. While the bases of the sinuses are well defined in the outline in Fig. 1, they are not clearly defined in those leaves where the lateral lobes are not well developed (Fig. 2). In order to have a consistent way of locating the two pseudolandmarks for all 86 leaves, they are defined as that point on the contour between the central lobe tip 'c' and a lateral lope tip 'r' or 'l', which is closest to the point of insertion of the petiole 's'.

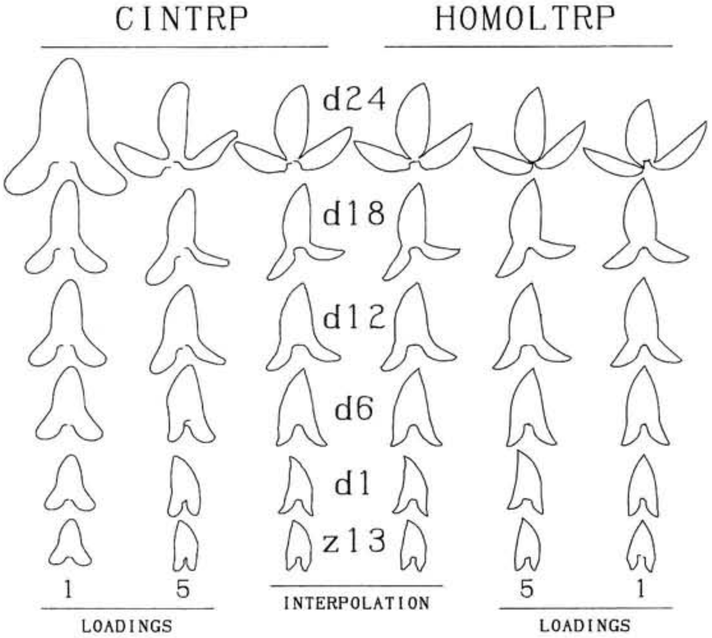

Fig. 2. A comparison of two eigenshape techniques, that based on a single homologous contour per outline (**cintrp**) on the left, and that based on six homologous contours per outline (**homoltrp**) on the right. The six rows of leaves (d24 to z13) represent six leaf outlines chosen to represent the full range of shape. Leaves d1 through d24 were gathered from a single shoot in which successive leaves were numbered consecutively, thus d24 was taken from the 23rd node above leaf d1. Leaf z13 was collected from a separate individual. The central two columns of leaves are the interpolations of 100 steps into the original digitized outlines, by the two techniques. The outer columns of leaves are the outlines reconstructed from the loadings on the first principal component (columns 1 and 6), and the outlines reconstructed from the loadings on the first five principal components (columns 2 and 5).

The six landmarks on the leaf outline yield six homologous contours: sr, ra, ac, cb, bl, and ls. Next, each contour is divided into some number of equal-length steps so that the contour can be represented by a series of angles. It is necessary that homologous contours on different specimens be divided into the same number of steps. Different contours on the same specimen may be represented by the same or a different number of steps. The number of steps chosen for any contour is arbitrary and could be adjusted in proportion to the complexity of the particular contour, to the length of the particular contour, or any other criteria.

In this example, the number of steps for each of the six contours was chosen to reflect the proportional length of that contour on the average outline (Fig. 1). In the average outline, the six contours sr, ra, ac, cb, bl, and ls have proportional lengths of 0.164, 0.117, 0.213, 0.234, 0.110, and 0.162, and they have been divided into the following number of steps: 16, 12, 22, 23, 11, and 16, respectively. This gives a total of 100 steps to represent the entire outline.

Within a given contour, steps will be interpolated into the contour such that all steps are equal in length. Each step can then be represented by the angle that the vector traced by that step makes in the plane. While the steps within a contour will be of equal length, the steps in different contours of the same specimen will generally be of different lengths. Therefore, it is necessary to preserve this relative length information. In order to keep the information on absolute size separate from the information on shape, the lengths of the contours will be recorded as proportional lengths: the length of a contour divided by the sum of the lengths of all contours (in this application all contours sum to the perimeter of the leaf outline and individual contour lengths are recorded as a proportion of the perimeter length).

As a result of this procedure, each outline is represented by 100 angle measures and six proportional length measures (one of the angle measures is redundant because the outlines are closed, and one of the proportional lengths is redundant because the perimeter sums to one). The number of angle measures used in any specific case is arbitrary and depends on how much resolution is desired. The number of contour length measures used will correspond to the number of homologous contours on each specimen. The number of contours will be determined by the number of landmarks.

The data for this study yielded 106 measures for each of 86 specimens. These data were arranged into a matrix which was normalized and then analyzed by singular value decomposition (Golub and Reinsch, 1970; Kennedy and Gentle, 1980; Press et al., 1986, 1988), which yields orthogonal eigenvectors and their associated eigenvalues. Successive eigenvectors represent successive principal components of variation in the data.

To the extent that many of the original variables (the 106 measures) covary, we may be able to capture and describe a large proportion of the variation in the shape with a relatively small number of variables (e.g., five or less). The objective of the technique is to reduce the variation in shape (as deduced from a large number of measurements) to a manageable number of variables that describe most of the variation. This reduced set of variables can then be used in systematic or other analyses as indices of shape. My application utilizes the reduced set of variables to characterize allometric changes in leaf shape during shoot ontogeny (Ray, 1990b).

Interpolation—The technique presented here is a modification of that developed by Lohmann and Schweitzer (1990), which involves the interpolation of some number of approximately equal-length steps into the entire outline. The method presented here divides the outline into a series of segments (the homologous contours) before the steps are interpolated. Step lengths are made equal only within each contour; however, the number of steps interpolated into a specific contour must be equal in every specimen. The interpolation algorithm of Schweitzer and Lohmann (written by C. Denham [WHOI], modified by Lohmann, recorded by Schweitzer) will be referred to as **cintrp**, and the interpolation algorithm used here will be referred to as **homoltrp**.

In order to represent a contour with a series of n interpolated steps, we must specify the length and direction of each step, resulting in a total of $2n$ variables. However, in eigenshape analysis we reduce the number of variables to $n + 1$ by making all of the steps equal in length. Now we need keep only the n angle measures and the one length.

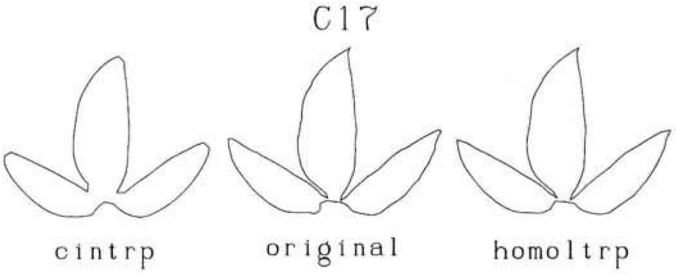

Fig. 3. A comparison of two interpolation methods. The central leaf (c17) is drawn from the 128 coordinates produced by the image analysis system. On the left (**cintrp**) is the outline resulting from the interpolation of 100 steps by the algorithm of Schweitzer and Lohmann. On the right (**homoltrp**) is the outline resulting from the interpolation of 100 steps by the algorithm described in this paper. In this example, both **cintrp** and **homoltrp** interpolated only one homologous contour for the entire outline, only one landmark ('s') was fixed.

In order to allow reconstruction of shapes from principal components, an interpolation algorithm must produce steps that are equal in length, with length measured as the straight line distance between the endpoints of the steps. The **cintrp** algorithm, however, does not measure, nor attempt to equalize, the straight line lengths of the steps. It equalizes the lengths as measured along the outline. To the extent that the outline is bent within a step, the linear distance will be shortened.

The **cintrp** algorithm attempts to reduce this effect somewhat by iteration; however, this only creates another problem, because successive iterations are not made against the original outline, but against the last iteration. This results in erosion inward from any bend, and is most severe in the vicinity of sharp curves where many landmarks are found. Figure 3 shows an original digitized leaf outline and the interpolation produced by the **cintrp** algorithm compared to that produced by the **homoltrp** algorithm. However, the erosive effects are not obvious for many outlines (Fig. 2).

The segmented nature of the interpolation is the most significant difference between the **homoltrp** procedure and that of Lohmann and Schweitzer (1990). However, **homoltrp** uses a completely different approach to the interpolation itself. Interpolation of an arbitrary number of equal-length steps into an arbitrarily complex outline is not a trivial problem. There is generally no exact solution, and continuous changes in step length result in discontinuous changes in the positions of steps. In order to deal with this bad behavior and produce nearly equal length steps, the **homoltrp** interpolation procedure involves a combination of three techniques. The first technique is identical to the first pass of the **cintrp** procedure, except that each homologous contour is treated separately, as **cintrp** would treat the entire outline:

First pass—We divide the total length, ol, of the original digitized contour by the number of steps, n, we wish to interpolate in order to calculate the interpolated step length $di = ol/n$. Measuring along the original digitized outline, we determine a series of interpolated points corresponding to the endpoints of the n steps of length di. While these steps will be equidistant along the original (bent) digitized outline, the straight line distance between successive interpolated points will generally be less than di, due to the bending of the original outline.

In the method of **cintrp**, the interpolation of the entire outline is repeated by replacing the original data with the interpolated data, then interpolating again, until the length of the interpolated contour is equal to the length of the contour into which it was interpolated to within a tolerance of 1%.

Second pass—The cumulative length of the interpolated contour il will be less than the length of the original contour ol. For the second pass, the step length is calculated as: $di = il/n$. This is the only information that we retain from the first pass; the actual interpolated points and steps are discarded.

On the second pass, step lengths are not measured along the original contour, but as straight line distances. This interpolation is performed successively for the first $n - 1$ steps, and these will all have exactly the desired length di. However, we are left with an n^{th} step, which will generally not be of the desired length. We may stop here, or test the deviation of the last step from the desired length. If it is not close enough to satisfy our criteria, we can calculate a new cumulative length for the interpolated contour and repeat the interpolation of the n steps.

The problem with iterating this interpolation algorithm is that it behaves badly for contours that have corners with angles of less than 90°. Also, geometric analysis shows that if the contour includes a process whose width is less than the step length, there may be no exact solution; that is, it may not be possible to interpolate exactly equal length steps. This situation may also hold for contours with corners. Observation shows that minor adjustment of the step length causes the interpolated outline to pop into or out of corners; in the vicinity of corners a continuous change in the step length causes a discontinuous change in the positions of the interpolated steps.

Third pass—As a way around the bad behavior of the rigid interpolation algorithm in the vicinity of corners, a procedure called **massage** has been developed. The massage routine examines every pair of adjacent steps on the contour and identifies the pair whose lengths are the most dissimilar. After the second pass, these will be steps n and $n - 1$. **Massage** moves the point joining the two most dissimilar steps, along the original contour, so that they will have equal lengths. It then finds the new most dissimilar pair of steps and makes their lengths equal. This process can be repeated until all the steps on the contour are equal to within the desired criterion, or a fixed number of times after which we perform another rigid interpolation of equal length steps.

In practice, ten iterations of the **massage** algorithm have been employed between each iteration of the rigid interpolation algorithm. However, some outlines never converge to the desired criteria (no two adjacent steps may differ in length by more than 0.1%), so iteration terminates after some fixed number of iterations. Although not perfect, the result is largely equal-length steps. **Homoltrp** produces a greater uniformity of step lengths than the **cintrp** algorithm, and because each iteration is made against the original outline, the interpolated outline cannot erode away from the original.

Working with angles—The interpolation finds a series of points that define equal-length steps, but these steps are converted into angles at the time they are output from the algorithm. The angle data will be fed into the singular value decomposition (SVD) routine. Before passing the data to the SVD routine, some normalization may be desired, and therefore we must come to grips with some of the special properties of angular data.

The first property to consider is that distinguishable angles only occupy a range of 2π. Beyond this range there is redundancy. Allowing angles to grow out of this range through additions or subtractions will cause corruption of the contours. This can be prevented by keeping the angular data within a range from $-\pi$ to π.

The next property of angles to consider is that while they can be safely added and subtracted, it is not feasible to calculate a mean through conventional methods. Consider the result of calculating the mean of the two angles π and $-\pi$, which actually are two ways of referring to the same angle. The desired result is either π or $-\pi$, but a simple numeric mean will give a result of 0. Mean angles can be computed through a vector sum procedure, in which each angle θ_i is represented as a unit vector in the direction θ. The unit vectors are summed, and the angle of the resultant vector is the mean angle. If the length of the resultant vector is divided by the number of vectors entered into the sum, the length of this normalized resultant vector is inversely proportional to the standard deviation of the vector sum, according to the formula: $s = [2 * (1 - r)]^{1/2}$, where s is the standard deviation and r is the normalized length of the resultant vector. The procedure for calculating angular means and standard deviations is detailed in Batschelet (1965) and Strong and Ray (1975).

Finally, in order that the angular measures made on each specimen are invariant to rotation of the specimen in the plane, it is desirable to root the angle measures for each specimen. This can be done by choosing some line segment on each specimen that can be considered standard, and using it to define the zero direction. Lohmann and Schweitzer (1990), following the procedure of Zahn and Roskies (1972) (see also Rohlf and Archie, 1984), rooted the angles on the first step of the outline by setting its direction to zero. This is a reasonable solution for the ostracode data set as they have no landmarks that could be used to define another line segment.

The problem with rooting by the angle of the first step in the more general case is that it makes the angular coordinate system for each specimen sensitive to the angle of the first step, which may vary widely in relation to the orientation of the specimen (as it does for the data set presented here). In order to compensate for this effect, it becomes necessary to use the unconventional normalization procedure of Lohmann and Schweitzer (1990), normalization by specimen. When the angle measures are normalized across a specimen, and the mean angle is subtracted from each angle measure, the result is a rotation of all specimens, bringing them into alignment, thereby compensating for the effects of the variable angle of the first step.

The method used for this data set is to root the angles on the line connecting landmark 's' to landmark 'c', the midrib of the leaf. This is considered to be the most stable line segment in the collection of specimens, and keeps the specimens in alignment to one another.

Normalization—The data were normalized in the traditional method: each of the 106 measures was summed across the 86 specimens in order to compute the mean and standard deviation. Arithmetic means and standard deviations were computed for the six proportional length measures, and angular means and standard deviations were calculated for the 100 angle measures. From each measure, the associated mean was subtracted, and the difference was divided by the associated standard deviation. This is in contrast to the normalization used by Lohmann and Schweitzer (1990), who summed across the angle measures to compute the mean and standard deviation for each specimen. The significance of Lohmann and Schweitzer's normalization procedure is discussed by Ray (1990a).

Singular value decomposition—The normalized matrix of 106 measures over 86 specimens was passed to the **svdcmp** algorithm published in Numerical Recipes in C (Press et al., 1988; the same algorithms are available in the Pascal and Fortran languages, Press et al., 1986). For the $M \times N$ data matrix $\mathbf{A}$, whose number of rows (specimens) M, is greater than or equal to its number of columns (measures) N, the singular value decomposition yields the following result: $\mathbf{A} = \mathbf{U}\mathbf{S}^{1/2}\mathbf{V}^{\mathrm{T}}$. $\mathbf{U}$ is an $M \times N$ matrix and $\mathbf{V}$ is an $N \times N$ matrix. Both are orthogonal, in the sense that the columns of $\mathbf{U}$ are orthonormal and both the columns and rows of $\mathbf{V}$ are orthonormal: $\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{V}\mathbf{V}^{\mathrm{T}} = \mathbf{1}$. The matrix $\mathbf{S}$ is an $N \times N$ diagonal matrix (the off diagonal elements are zero) whose entries are the eigenvalues of the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$. These eigenvalues are equal to the variance accounted for by each of the eigenvectors (Kennedy and Gentle, 1980). The eigenvectors represent a least squares fit to the principal components of variation of the data (Press et al., 1988).

The proportion of the variation accounted for by each eigenvector can be calculated by first computing the total variance, the sum of the diagonal elements of $\mathbf{S}$. For each eigenvector, the associated value of $\mathbf{S}$ is divided by the sum to get the proportion of the total variance accounted for by that principal component.

Given that the algorithm requires that the data matrix have at least as many rows as columns, it is worth noting that given the above relations, the following also holds: $\mathbf{A}^{\mathrm{T}} = \mathbf{V}\mathbf{S}^{1/2}\mathbf{U}^{\mathrm{T}}$. The data set presented here has more columns than rows. Press et al. (1988) recommended augmenting the matrix with rows of zero in this case. However, the computation can be speeded and effected more cleanly by taking the transpose of $\mathbf{A}$ before the decomposition and rearranging the results accordingly.

As a result of these operations, the matrix $\mathbf{U}^{\mathrm{T}}$ contains the 86 eigenshapes (principal components), each with 106 values. The dot product $\mathbf{V}\mathbf{S}^{1/2}$ results in the 86×86 matrix of loadings $\mathbf{L}$. The dot product $\mathbf{L}\mathbf{U}^{\mathrm{T}}$ gives back the original shapes, the matrix $\mathbf{A}$. However, the original shapes can be approximated by taking the dot product using only the first few loadings. In this way, the number of variables required to describe the variation in shapes can be greatly reduced.

It took the **svdcmp** algorithm 16 minutes to process the matrix when running on the Zenith 183, which uses an 8 mhz 8088 processor and an 8 mhz 8087 math co-processor. The same program took 2 minutes on the Toshiba 5200 which uses a 20 mhz 80386 processor and a 20 mhz 80387 math co-processor. The **svdcmp** algorithm is called from a larger program that normalizes the data and synthesizes outlines from the eigenshapes produced by the **svdcmp** program. The program has a memory-intensive windowing interface, yet at its most loaded it used less than 260 kb of memory. The program may be obtained from the author. It accepts as input any ordered list of x, y coordinates making up a closed outline.

RESULTS

The principal objective of the morphometric technique presented here is to describe the greatest amount of the variation in the data with the smallest number of variables. In this section a comparison is made of the results of eigenanalysis of the data derived by the application of the **cintrp** algorithm to interpolate a single homologous contour into each specimen, with the results of eigenanalysis of data derived by the application of the **homoltrp** algorithm to interpolate six homologous contours into each specimen. The former will be referred to as the "**cintrp**" or "eigenshape" technique and the latter as the "**homoltrp**" or "landmark eigenshape" technique.

The efficiency of techniques using eigenanalysis can be compared by utilizing the fact that the sum of all eigenvalues from the analysis is equal to the total variance of the data, and the proportion of the variance explained by any single factor (eigenvector) is just its eigenvalue divided by the total variance. Thus, for example, we can compare what proportion of the variance is accounted for by the first principal component, or by the first five principal components.

Using this approach, we find that the first five principal components from the **cintrp** technique account for 0.715, 0.115, 0.036, 0.023, and 0.021 of the total variance of the data, for a cumulative 0.910 of the variation. The first five principal components from the **homoltrp** technique account for 0.487, 0.187, 0.042, 0.040, and 0.037 of the total variance, for a cumulative 0.793 of the variation. This comparison suggests that the **cintrp** technique is much more efficient. However, a visual comparison of the reconstructions of leaf outlines from the first and the first five principal components (Fig. 2) suggests that the landmark eigenshape technique is more efficient.

These conflicting observations can be resolved in part by noting that the two techniques measure variation from different bases. In the **cintrp** technique, angles on outlines are recorded as deviations from the corresponding angles on a circle. In the **homoltrp** technique angles are measured as deviations from the corresponding angle on the average outline, Fig. 1. In the **cintrp** technique the average outline is located not by the normalization process, but by the first principal component. Due to these fundamental differences in the techniques, their relative efficiencies can-

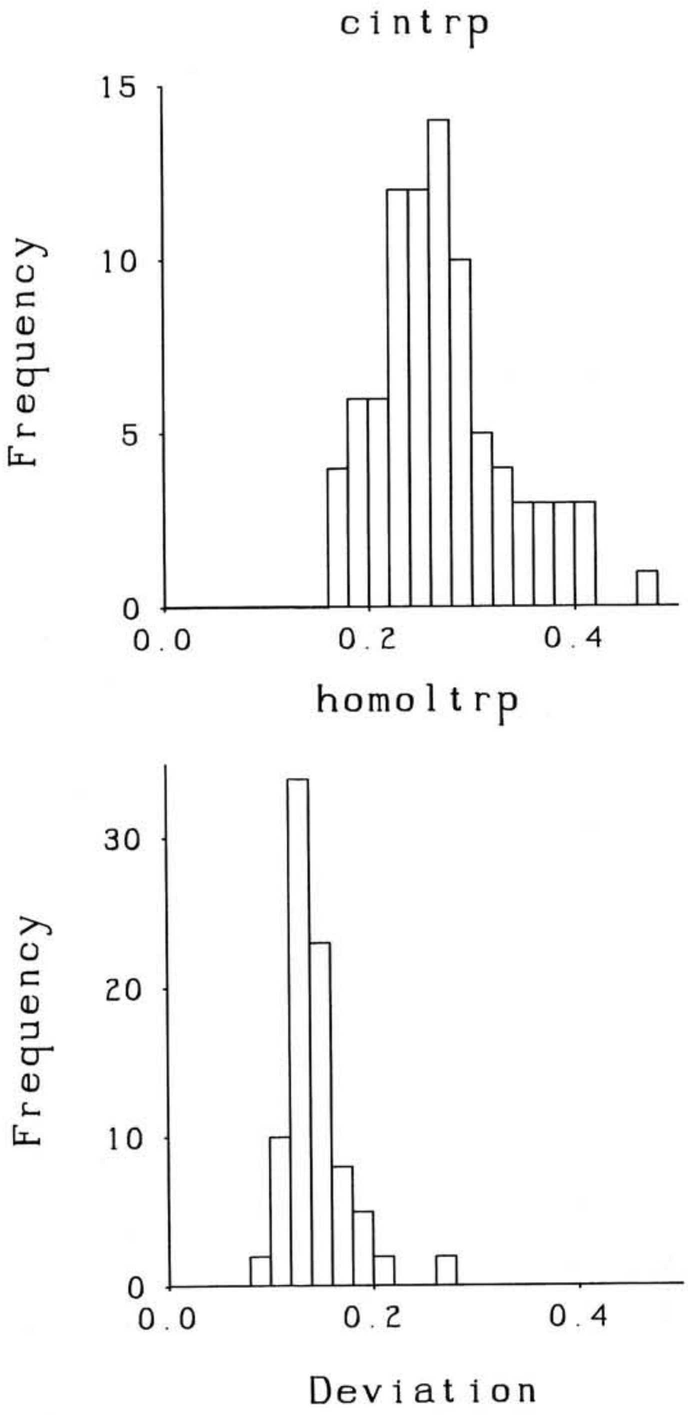

Fig. 4. The distributions of the deviations between the 86 original interpolated outlines and the corresponding outlines reconstructed from the loadings on the first five principal components. The method based on a single homologous contour (**cintrp**) above and the method based on six homologous contours (**homoltrp**) below.

not be evaluated by comparing the proportion of the variance explained by the first *X* principal components.

In order to provide a more appropriate comparison, an ad hoc statistic has been developed that reflects the deviation between the angles of a reconstructed outline and those of the original outline. Based on the formula for the standard deviation, the deviation between outlines is calculated as follows:

$$d = \left[\sum_{i=1}^{n} (\theta o_i - \theta s_i)^2/(n - 1) \right]^{1/2}$$

where θo_i is the i^{th} angle on the original interpolated outline, and θs_i is the i^{th} angle on the synthesized outline. A value of the deviation can be calculated for each leaf, providing a distribution of 86 deviations for each reconstruction of the complete set of leaves by either technique. These distributions can be compared directly to determine how changes in the technique affect the efficiency of the reconstruction.

Figure 4 compares the distributions of deviations from the **cintrp** technique to those from the **homoltrp** technique. The **homoltrp** technique yields significantly lower deviations. In fact, the outlines reconstructed from only the loadings on the first principal component from the **homoltrp** technique result in (insignificantly) less deviation than the reconstructions from the first five principal components from the **cintrp** technique (Table 1; Fig. 2).

The landmark eigenshape (**homoltrp**) technique as applied here differs from the eigenshape (**cintrp**) technique as applied here in four ways:

Landmark Eigenshape Analysis, homoltrp—

1) Interpolation is based on six homologous contours per outline, lengths of steps are equal within contours, homologous contours are interpolated with equal numbers of steps.
2) Data are normalized by measures (for each angle or length measure, sum across all specimens to compute the mean and standard deviation).
3) Outline is represented by a ϕ function (circle not removed).
4) Angles are rooted on the midrib of the leaf.

Eigenshape Analysis, cintrp—

1) Interpolation of outlines uses one landmark, all steps within an outline are equal in length (one homologous contour per outline).
2) Data are normalized by specimens (for each specimen, sum across all angle measures to compute the mean and standard deviation).
3) Outline is represented by a ϕ^* function (circle removed).
4) Angles are rooted on the first angle of the outline.

In order to demonstrate that the decrease in deviations is due primarily to the use of homologous contours, and not the other differences, some additional combinations of these differences have been compared and the results are presented in Table 1. The table also includes an example of interpolation of 94 steps by the **homoltrp** method. This is to address the possibility that perhaps the extra efficiency of this method derives from having more information available, 106 total measures as opposed to the 100 total measures of the **cintrp** method. The results do not support this idea.

TABLE 1. *Comparison of eigenshape analysis techniques*

Interpolation[a]	Outline[b]	Normalization[c]	Rooting[d]	Loadings[e]	Mean[f]	SD[g]
C 100	ϕ*	S	F	5	0.272	0.064
C 100	ϕ	M	F	5	0.267	0.059
C 100	ϕ	M	M	5	0.262	0.057
H 100, 6	ϕ	M	M	1	0.264	0.078
H 100, 6	ϕ	M	M	5	0.145	0.030
H 94, 6	ϕ	M	M	5	0.144	0.030

[a] C 100 = **Cintrp** style interpolation of 100 angles into each outline. H 100, 6 = **Homoltrp** style interpolation of 100 total angles, distributed within six contours in each outline. H 94, 6 = **Homoltrp** style interpolation of 94 total angles, distributed within six contours in each outline.

[b] ϕ* = circle removed, angle measures represent deviation of each angle from corresponding angle on a circle. ϕ = circle not removed, angle measures represent actual angular direction in the plane.

[c] S = normalization by specimen. M = normalization by measures.

[d] F = rooting angles by setting the first angle on the outline equal to zero. M = rooting angles by setting the line from landmark 's' to landmark 'c' (the midrib of the leaf) equal to zero.

[e] 1 = only the loading on the first principal component is used in constructing the synthesized outline. 5 = the loadings on the first five principal components are used in constructing the synthesized outline.

[f] The mean computed for the 86 leaves of the sample, of the measure of deviation $\left\{d = \left[\sum_{i=1}^{n} (\theta o_i - \theta s_i)^2/(n-1)\right]^{1/2}\right\}$ between the original interpolated outline and the outline synthesized from the loadings on the first principal component(s).

[g] The standard deviation of the 86 measures of deviation.

DISCUSSION

The work presented here shows that the efficiency of eigenanalysis of outlines can be improved by segmenting the outline between homologous landmarks. This is not a surprising result and is quite easily accomplished. The technique was applied here to closed outlines, as "outline" morphometric techniques generally are. However, recognition of the concept of homologous contours opens the way to the application of eigenanalysis to sets of contours other than closed outlines.

There is a wide set of curves that could be usefully analyzed, that either are not closed or are not connected. These might include (homologous) portions of outlines, or the outlines of appendages such as anthers, stigmas, or petals. In addition, there are contours that are not parts of the outline, but occur internally on the specimen, such as the contours of the principal veins of leaves, or suture lines or ridges on seeds, fruit, or pollen.

One might want to treat a collection of contours that are not all connected. In these cases there is need for additional information to place the individual contours relative to one another. When contours do not share common endpoints, their relative positions could be determined by setting up a truss of the bounding landmarks (Strauss and Bookstein, 1982; Bookstein et al., 1985; Dickinson, Parker, and Strauss, 1987).

Use of a truss to define the relative positions of the contours suggests an alternative means of rooting the angles of the contour steps and determining the proportional lengths of the contours. For each individual contour, the line connecting the two boundary points could define the zero direction and the unit length. While each contour would then have a distinct root direction, the relationships between the individual root directions would be defined by the truss. Similarly, the length of the contour or of the individual steps could be expressed in relation to the distance between the endpoints. Step lengths could be expressed as a proportion of the straight line distance between the endpoints of the contour. The truss system would determine the relationships between the actual lengths of the different contours. To make this idea clear and concrete, I suggest the C data structure of Table 2 for the contour data.

Using this approach would make the individual contours into self-contained data objects that could be entered into a truss system along with landmarks not associated with contours, providing a complete integration of the landmark and contour techniques. In view of these possible approaches, and the success of the present application, the method warrants further experimentation. In addition, it would be valuable to explore the sensitivity of the technique to the choice of the number of steps used for each contour and to compare the technique to other "outline" techniques that could be similarly segmented.

Landmark eigenshape analysis, the new technique presented here, is one of a growing number of morphometric techniques, which, combined with new automated data

TABLE 2. *Suggested C data structure for contour data*

```
struct point {float x; float y;}; /* x and y coordinates of point */
struct landmark {
  char        n[ ]; /* name of landmark */
  struct point  p;     /* x and y coordinates of landmark */
};
struct contour {
  char          n[ ]; /* name of contour */
  struct landmark b;   /* beginning point of contour */
  struct landmark e;   /* end point of contour */
  int s;   /* number of equal length steps on contour */
  float l;   /* proportional length of step:
                (step length) / (distance between landmarks b and e) */
  float a[ ]; /* array of s angles: the angular direction of each step,
                relative to the direction from landmark b to landmark e */
  float d[ ]; /* array of s lengths: the length of each step relative to l,
                this is an accuracy check of the interpolation algorithm */
  };
```

acquisition technology, allows an increasingly sophisticated treatment of shape. With these new techniques, shapes can be quantified in ways that allow their comparison or the study of their changes. Comparisons would generally be made on the basis of the first few (five or less) principal components of shape. We may use these techniques to determine the degree of similarity of leaf shapes in systematic studies. We may characterize the gradual changes of shape during development of leaves, flowers, or seeds, or the heteroblastic changes of leaves or other organs along shoots. We may develop more powerful means of characterizing or testing for differences in shapes due to ecological factors, such as sun and shade leaves.

Landmark eigenshape analysis is a relatively sophisticated technique that allows the characterization of very subtle and difficult to describe differences in shape. Such a powerful technique cannot be recommended in all applications. However, where there is a need to capture a great deal of detail about shape with a few variables, landmark eigenshape analysis should be considered.

LITERATURE CITED

BATSCHELET, E. 1965. Statistical methods for the analysis of problems in animal orientation and certain biological rhythms. American Institute of Biological Sciences, Washington DC.

BOOKSTEIN, F. L. 1978. The measurement of biological shape and shape change, 191. Springer-Verlag, Berlin.

———, B. CHERNOFF, R. ELDER, J. HUMPHRIES, G. SMITH, AND R. STRAUSS. 1985. Morphometrics in evolutionary biology. Special Publication 15, 277. The Academy of Natural Sciences of Philadelphia, Philadelphia, PA.

———, R. E. STRAUSS, J. M. HUMPHRIES, B. CHERNOFF, R. L. ELDER, AND G. R. SMITH. 1982. A comment upon the uses of fourier methods in systematics. *Systematic Zoology* 31: 85–92.

DICKINSON, T. A., W. H. PARKER, AND R. E. STRAUSS. 1987. Another approach to leaf shape comparisons. *Taxon* 36: 1–20.

GOLUB, G. H., AND C. REINSCH. 1970. Singular value decomposition and least squares solutions. *Numerical Mathematics* 14: 403–420.

JENSEN, R. J. 1990. Detecting shape variation in oak leaf morphology: a comparison of rotational-fit methods. *American Journal of Botany* 77: 1279–1293.

KENNEDY, W. J., AND J. E. GENTLE. 1980. Statistical computing, 591. Marcel Dekker, New York, NY.

LOHMANN, G. P. 1983. Eigenshape analysis of microfossils: a general morphometric procedure for describing changes in shape. *Mathematical Geology* 15: 659–672.

———, AND P. N. SCHWEITZER. 1990. On eigenshape analysis. *In* F. J. Rohlf and F. Bookstein [eds.], Proceedings of the Michigan Morphometrics Workshop, 147–166. University of Michigan Museum of Zoology, Ann Arbor. Available from: Director, Museum of Zoology, The University of Michigan, Ann Arbor, MI 48109–1079.

MCLELLAN, T. 1990. Development of differences in leaf shape in *Begonia dregei* (Begoniaceae). *American Journal of Botany* 77: 323–337.

PRESS, W. H., B. P. FLANNERY, S. A. TEUKOLSKY, AND W. T. VETTERLING. 1986. Numerical recipes; the art of scientific computing, 818. Cambridge University Press, Cambridge.

———, ———, ———, AND ———. 1988. Numerical recipes in C: the art of scientific computing, 735. Cambridge University Press, Cambridge.

RAY, T. S. 1990a. Application of "eigenshape" analysis to second order leaf shape ontogeny in *Syngonium podophyllum* (Araceae). *In* F. J. Rohlf and F. Bookstein [eds.], Proceedings of the Michigan Morphometrics Workshop, 201–213. University of Michigan Museum of Zoology, Ann Arbor, MI.

———. 1990b. Metamorphosis in the Araceae. *American Journal of Botany* 77: 1599–1609.

ROHLF, F. J., AND J. W. ARCHIE. 1984. A comparison of fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Zoology* 33: 302–317.

SCHWEITZER, P. N., R. L. KAESLER, AND G. P. LOHMANN. 1986. Ontogeny and heterochrony in ostracode *Cavellina* Coryell. *Paleobiology* 12: 290–301.

STRAUSS, R. E., AND F. L. BOOKSTEIN. 1982. The truss: body form reconstruction in morphometrics. *Systematic Zoology* 31: 113–135.

STRONG, D. R., AND T. S. RAY. 1975. Host tree location behavior of a tropical vine (*Monstera gigantea*) by skototropism. *Science* 190: 804–806.

ZAHN, C. T., AND R. Z. ROSKIES. 1972. Fourier descriptors for plane closed curves. *IEEE Transactions on Computers* C21: 269–281.