Syst. Biol. 47(1):147–158, 1998

On Applications of Geometric Morphometrics to Studies of Ontogeny and Phylogeny

F. JAMES ROHLF

Department of Ecology and Evolution, State University of New York, Stony Brook, New York 11794-5245, USA. E-mail: rohlf@life.bio.sunysb.edu

The field of geometric morphometrics is relatively new (see Bookstein, 1991; Rohlf and Marcus, 1993) and has shown very rapid progress over the last few years. As might be expected during a period of rapid development, there can be technical problems in some of the pioneering studies as biologists attempt to apply the new tools. It is only now becoming clear how the new techniques should be combined in order to carry out comprehensive analyses of real data sets. Bookstein (1996b) gave a list of recommendations for such applications, and Bookstein (1996a) gave several comprehensive examples of morphometric analyses. However, these accounts do not address some of the types of applications that are of particular interest in systematic biology. Bookstein (1994) pointed out problems with using geometric morphometric methods in the usual character-based cladistic studies. He emphasized that morphometrics cannot supply homologous shape characters. The purpose of this note is to comment about some recent applications of morphometric methods in systematic biology.

This note is particularly concerned with the morphometric methods used by Zelditch et al. (1992, 1993, 1995), Swiderski (1993), Fink and Zelditch (1995), and Zelditch and Fink (1995). These methods were also used by Burke et al. (1996). For convenience, these studies will be referred to using the acronym Z&F. In these studies the authors investigated the use of partial warps (Bookstein, 1991) as variables in ontogenetic and taxonomic studies. They were impressed by their observation that differences in partial warps scores (Rohlf, 1993b) corresponded to shape differences that could be localized on the bodies of the organisms. They also found differences in these variables between developmental stages and between species. They concluded that partial warps could be interpreted and used as traditional taxonomic characters and would be useful in evolutionary studies.

Lynch et al. (1996) tried to interpret the partial warps they obtained in their study but they were cautious about using them in the ways advocated by Z&F. They suggested that extensive simulation studies needed to be done to validate the Z&F approach. Naylor (1996) investigated Z&F's approach using data based on a simulated phylogeny. Even though his simulated phylogeny was based on a sequence of simple morphological changes and had no homoplasy, the results showed high levels of homoplasy and did not recover the morphological changes used to create the phylogeny.

The present paper is concerned with theoretical problems rather than empirical problems requiring validation through simulations. The fundamental problem is that Z&F interpreted the partial warps as homologous and biologically meaningful variables—rather than as mathematically elegant but biologically arbitrary variables whose definition is not based on any covariance patterns in the data. There is also a problem in their choice of the so-called "reference." This problem is discussed first as it provides the background that makes it easier to explain the other, more important, problems. These include ignoring the fact that partial warps are usually highly correlated, using methods whose results are very sensitive to different choices of the reference, and a warp-by-warp examination and interpretation of shape variables. Some suggestions are also made for the ways in which morphometric methods can be used to study ontogeny and phylogeny. The companion note by Adams and Rosenberg (1998) examines the Z&F protocol in more detail to demonstrate the consequences of different choices of a reference and different rotations of the tangent space.

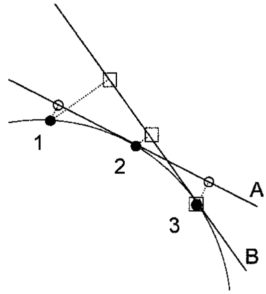

FIGURE 1. Projections of three points (●) in shape space onto two possible tangent lines. Tangent line A through the central point gives a less distorted representation of the relative positions of the projected points (○) than does tangent line B. On tangent line B, the projections (□) of points 1 and 2 are slightly closer together than the projections of points 2 and 3.

SHAPE SPACES AND THE CHOICE OF A REFERENCE

Landmark-based methods of geometric morphometrics are based on Kendall's (1981, 1984) definition of a shape space. Shapes (as captured by configurations of digitized landmarks) can be considered as points in this multidimensional space. The distances between points in this space are invariant to variation in the location, orientation, and scale of the coordinate system in which the specimens are digitized. It is assumed that there are a fixed number of landmarks, p. Although the geometry of Kendall's shape space is complicated, one can visualize some aspects as the $(kp - k - (k - 1)/2 - 1)$-dimensional surface of a high-dimensional sphere. The appropriate distances between points are Procrustes distances, which are geodesic (great circle) distances usually measured in radians. By definition, similar shapes are those that are close together and dissimilar shapes are far apart with respect to this metric. The power of the geometric approach is that this shape space captures all possible variations in shape of configurations of landmarks. It also excludes information about variation in translation, rotation, and size (sometimes these are called nuisance parameters because they are not of interest as descriptors of shape variation).

Because Kendall's shape space is a curved surface it has a non-Euclidean geometry. As a result conventional linear multivariate statistical methods are not appropriate. However, when variation in shape is relatively small it is possible to approximate the distance relationships between points by a linear space tangent to shape space just as a flat map can accurately approximate a small region of the earth's surface. This tangent space is of the same dimension as shape space. For example, in the case of shapes consisting of $p = 3$ landmark points digitized in $k = 2$ dimensions, one can visualize shape space as the surface of an ordinary sphere and the tangent space as an ordinary Euclidean plane touching its surface. Somewhat inappropriately, the tangent point is called the ''reference'' in geometric morphometrics. The statistical distribution of points in shape space is approximated by the distribution of the projections of the points from the surface of the sphere onto the tangent space (Fig. 1 shows a side view). Due to the foreshortening of the projections as one moves away from the reference (and shape space curves away from the tangent space), Euclidean distances between some pairs of points in tangent space are smaller than their corresponding Procrustes distances. This distortion becomes larger as one considers points further from the reference. The approximation is best when the reference is close to all of the points being projected. Using an extreme point as a reference would clearly be undesirable. Bookstein (1996b:145) recommended that one ''use a sensible reference form . . . best taken as the grand mean form'' to optimize the approximation of shape space by the tangent space. For this reason, the reference should be taken as the mean of the observed shapes after differences due to location, orientation, and size are removed using generalized orthogonal least-squares Procrustes analysis (GLS, Rohlf and Slice, 1990).

Zelditch et al. (1992), in a study of cotton rats from 1 to 30 days old, used the mean 1-day-old form as the reference (which they called the ''starting form''). Fink and Zelditch (1995), Zelditch and Fink (1995), and Zelditch et al. (1995), in studies of piranhas, used an average of three juveniles of an outgroup species as the reference. These choices were made because it seemed logical to compare different developmental stages to an initial one or to compare taxa to an outgroup. However, as discussed already, the reference in geometric morphometrics is not a standard to which other shapes are to be compared. The reference is simply the shape that defines the tangent space approximation so that linear statistical methods can be used to study shape. The choice of reference has nothing to do with the statistical comparisons to be made within the tangent space. To avoid this confusion about the purpose of the reference, Rohlf et al. (1996) referred to the reference as the ''tangent point'' or the ''tangent configuration.'' Once a tangent space is defined, one has a set of shape variables that can be used to make any desired comparisons. Developmental stages can be compared to an early stage or taxa to outgroups and so forth using standard statistical designs and multivariate methods within the tangent space.

A program, tpsSmall (Rohlf, 1997), is available to help one assess the accuracy of the approximation of shape space by the tangent space. It measures the accuracy by the correlation between Euclidean distances in the tangent space and Procrustes distances in shape space. Uncentered correlations are computed by the program because any line corresponding to the relationship between these distances must go through the origin. The approximation is usually very good.

The effects of different choices of a reference on distances in the tangent space are expected to be relatively minor. The well-known rat growth data (data from Bookstein, 1991) can be used as an example. It consists of eight well-spaced landmarks digitized around the brain case of 164 rats from 7 to 150 days of age. The effects of several different choices for the reference were investigated (the overall mean, the mean of the 7-day-old rats, and each of the first three 7-day-old specimens). The entire set of 164 specimens was projected onto tangent spaces defined by the different references by using matrices of partial warp scores based on each reference. These matrices were then used to compute matrices of distances among the 164 specimens. Uncentered correlations among the distance matrices for the three specimens and between each of them and the matrix based on the mean of the 7-day-old rats ranged from 0.99978 to 0.99998. The correlation between the distance matrix using the mean of the 7-day-old rats as a reference and the distance matrix using the mean of the entire data set as a reference was 0.99809. The partial warps ignore uniform shape differences (affine differences), which are very large in this data set. If Bookstein's (1996c) estimates of the uniform component are appended as additional variables the correlation rises to 0.99986. Thus the patterns of similarities and differences captured by the tangent space are quite stable; that is, the relative distances between the points in the tangent space are not changed much due to changes in the reference. Even though it may make little difference, the GLS consensus should be used as the reference because it is easy to compute and it minimizes errors in the approximation of shape space. On the other hand, the individual partial warps themselves are not very stable. This problem is discussed in the next section.

Shape Variables

When one uses conventional statistical analyses to study variation, to compare shapes, or to explore the covariation of shape with extrinsic variables, one needs values for each specimen on a set of shape variables. Shape variables can be generated in many different ways. The choice is important because different sets of variables have different interpretations and may lead to different statistical conclusions in some kinds of analyses. A simple set of shape variables is the set of kp coordinates of the p landmarks in k dimensions for n specimens after they have been aligned to the reference using GLS superimposition. The coordinates can be interpreted as variables that, when taken together, define the position of each specimen in tangent space. However, these variables are redundant because tangent space has only $kp - k - k(k - 1)/2 - 1$ dimensions (after the removal of the effects of differences in translation, rotation, and size). One can obtain a nonredundant set of shape variables by using Bookstein shape coordinates (Bookstein, 1991) to align the specimens along one or more baselines. The resulting shape variables are easy to understand, but they do not yield as good an approximation of shape space as do the residuals from a GLS superimposition.

Another method (the one used by Z&F) is to use Bookstein's (1989) partial warps as shape variables. See Bookstein (1989, 1991) or Rohlf (1996) for the details of their computation and interpretation. Only the properties relevant to the present discussion are presented here. The $p - k - 1$ principal warps are eigenvectors of a bending energy matrix and are computed using only information from the reference configuration. The principal warp vectors allow one to describe the ways in which it is possible to deform the shape of the reference configuration. The partial warps are computed by projecting the x and y (and possibly z) coordinates of the aligned specimens separately onto each of the principal warp vectors. The partial warps are shape variables that can be viewed as geometrically orthogonal coordinate axes for the part of tangent space that corresponds to shapes that can be described in terms of nonlinear deformations (localized changes) of the reference configuration. This is often called the nonaffine shape component. One can also generate shape variables that describe shapes that correspond to affine (i.e., nonlocal) deformations of the reference (see Bookstein, 1996c). These are usually referred to as the uniform shape component and can be thought of as the zeroth partial warps.

Partial warps have the mathematically elegant interpretation of giving a coordinate system for the tangent space that also corresponds to a decomposition of the potential modes of shape variation into geometrically orthogonal components at different spatial scales. Principal warps with large eigenvalues (large bending energies) correspond to small-scale shape changes, and those with small eigenvalues correspond to large-scale shape changes. Note that the patterns of shape changes corresponding to each warp are based only on the configuration of landmarks in the reference (Bookstein, 1991). No information on shape variation or covariation is taken into consideration (Rohlf, 1993b, 1996). The partial warp scores for a specimen are scalar values for each dimension indicating how much of each principal warp is needed to account for the observed differences between the specimen and the reference. The specimens have no influence on the kind of shape change corresponding to each principal warp and hence have no influences on the meaning of each partial warp.

Although the tangent space is quite stable (as discussed earlier), the orientations and hence the interpretations of the partial warp axes are not very stable. One way to demonstrate this is to compute angles between principal warps based on different references. The angles can be computed as the absolute value of the arccosines of the product of one matrix of normalized principal warps times the transpose of another such matrix. Ideally, the result would be a matrix with angles of 0° down the diagonal and angles of 90° in the off-diagonal entries. The entries down the diagonal are shown in Table 1 for the rat data. The angles for comparisons of the first three specimens with the average of the 7-day-old rats are given in the first three rows. The angles varied from 24.8° to 87.8°. The com-

TABLE 1. Angles (in degrees) between principal warps based on different references for the rat data from Bookstein (1991). Warps are ordered from smallest to largest spatial scales.

Comparisons	Principal warps 1	2	3	4	5
Specimen 1 vs. mean 7 day old	33.7	63.0	69.6	74.0	87.8
Specimen 2 vs. mean 7 day old	38.6	39.0	39.8	44.5	54.8
Specimen 3 vs. mean 7 day old	73.1	45.9	24.8	51.7	84.1
Mean 7 day old vs. overall mean	89.0	87.9	70.1	64.5	72.4

parison of the principal warps using the average of the 7-day-old rats as a reference versus using the mean of the entire data set as a reference are given in the last row. The angles varied from 64.5° to 89.0°. There was also high variability in the off-diagonal angles in each of the comparisons. Thus the pattern of shape changes implied by each principal warp (and hence the biological interpretation of each partial warp) is very sensitive to the choice for the reference. One expects the smallest scale (most localized) principal warps to be quite sensitive to small changes in the shape of the reference configuration. Another way to demonstrate the lack of stability is simply to move a landmark in the reference and visualize the new principal warps as thin-plate splines (an example using artificial data is given later).

Zelditch et al. (1995) emphasize that the partial warps corresponded to shape differences that can be localized on the body of the organisms. This is not, however, a unique property of partial warps. The entire space spanned by the partial warps (other than the 0th) as well as all linear combinations of the partial warps have this property. As Zelditch et al. (1992:1169) pointed out, "There is no biological information . . . in the coefficients for each landmark in these eigenvectors. Rather, these principal warps provide a basis for comparison, a list of features, each progressively more localized, for comparisons of differences between forms." This statement also applies to the partial warps because the partial warps are just the principal warps multiplied by scalar weights for each dimension. One expects the space spanned by the partial warps to contain useful and biologically interpretable information—they span the part of tangent space that contains information on all possible shape variables that can be localized. As discussed in the next section, there is no reason to expect the partial warps to represent directions that should be of particular biological interest. Even though the landmarks may be homologous, it is difficult to think of arbitrary linear combinations as being homologous. Curiously, Zelditch et al. (1992) described and interpreted each partial warp and performed tests of significance for age differences for each partial warp separately.

BIOLOGICAL INTERPRETATION OF PARTIAL WARPS?

An important problem with the Z&F studies is that they use and interpret each partial warp separately. Zelditch and Fink (1995:343) correctly observed that a limitation of the interpretation of Bookstein shape coordinates for studies of developmental integration is that "we cannot know in advance of analysis which triangles span developmentally autonomous units and which combinations of triangles extend over developmentally integrated areas." Partial warps have the same kind of limitation because they are simply weighted principal warps whose pattern of deformation is determined in advance by the configuration of landmarks in the reference and not by the patterns of covariation in the data.

Zelditch et al. (1992:1176) reported a higher level of integration of the different parts of the skull than they expected based on conventional interpretations of mammalian skull growth. They also reported a lack of evidence for developmental units corresponding to the facial skull or cranial base. Their conclusions were based on their observation that the large-scale partial warps described shape deformations across many different parts of the skull and that the effects of the small-scale partial warps were not localized corresponding to the traditional developmental units. These conclusions are unwarranted. The deformations corresponding to the principal warps (and hence the partial warps) are, as discussed earlier, functions of just the configuration of landmarks in the reference and does not take into consideration patterns of covariation among the landmarks. The fact that a partial warp does or does not align with biologically meaningful structures is fortuitous.

The following artificial example may be helpful. Figure 2a shows a configuration with three landmarks in each of two groups. Using this configuration as a reference, the principal warps can be computed (they are shown separately as x and y deformations). Figure 3 shows a configuration of landmarks in which the upper group of landmarks has been expanded to represent a simple change in a single biological feature. This change can also be shown as a thin-plate spline deformation of the reference as shown in Figure 3b. Figure 4 shows the decomposition of the deformation in Figure 3b into its three partial warps. The x and y contributions to each partial warp are shown jointly. Partial warp 1 is principal warp 1 with a weight of 0.0 for the x-axis and −0.46 for the y-axis. Because the coefficient is negative for the y-axis, the implied deformation is a relative expansion of the sizes of both groups of landmarks. The second partial warp is principal warp 2 with a weight of 0.0 for the x-axis and 0.58 for the y-axis. This expansion of the upper group of landmarks relative to the lower group is the type of pattern that one would intuitively expect. However, it is too extreme in its compression of the lower region. The third partial warp is the third principal warp with weights of 1.0 for the x-axis and 0.0 for the y-axis. This warp also shows an extreme compression of the lower group of landmarks but this time in the x direction. The second and third partial warps are extreme in order to compensate for the pattern in the first partial warp that shows an expansion in the lower group of landmarks. Note that in a sense the effects of the partial warps are not really localized—their effects are best described as contrasts in which the expansions of some regions are relative to compressions in others. The partial warps in Figure 4 might seem to suggest integration because the warps have a complementary effect on the two regions. This is an unwarranted conclusion because the partial warps cannot show any pattern of deformation that is not given by the principal warps, and the principal warps are only a function of the reference, not of the covariance among the landmarks. Separate warp-by-warp examinations of the partial warps cannot reveal either a pattern of partitioning of landmarks into covarying groups nor their integration across groups.

Note that three partial warps were required to yield the simple morphological difference that was used to create the configuration in Figure 3a. Because the warps are based only on the reference, it is unlikely that one of them would correspond to an appropriate morphological change. Similar results were found by Naylor (1996) in his simulations. Figure 5 demonstrates the sensitivity of the meaning of a warp to small changes in the reference. Figure 5a differs from Figure 2a only in that landmark 2 has been moved closer to landmark 1. Figures 5b and 5c show the resultant smallest scale principal warp in terms of x- and y-deformations. The meaning of this warp has changed considerably from that shown in Figures 2b and 2e. In view of these properties of principal and partial warps, it does not make much sense to try to interpret them as homologous characters. Bookstein (1994) discusses several additional theoretical reasons why partial warps should not be treated as homologous characters. However, when taken together the partial warps de-

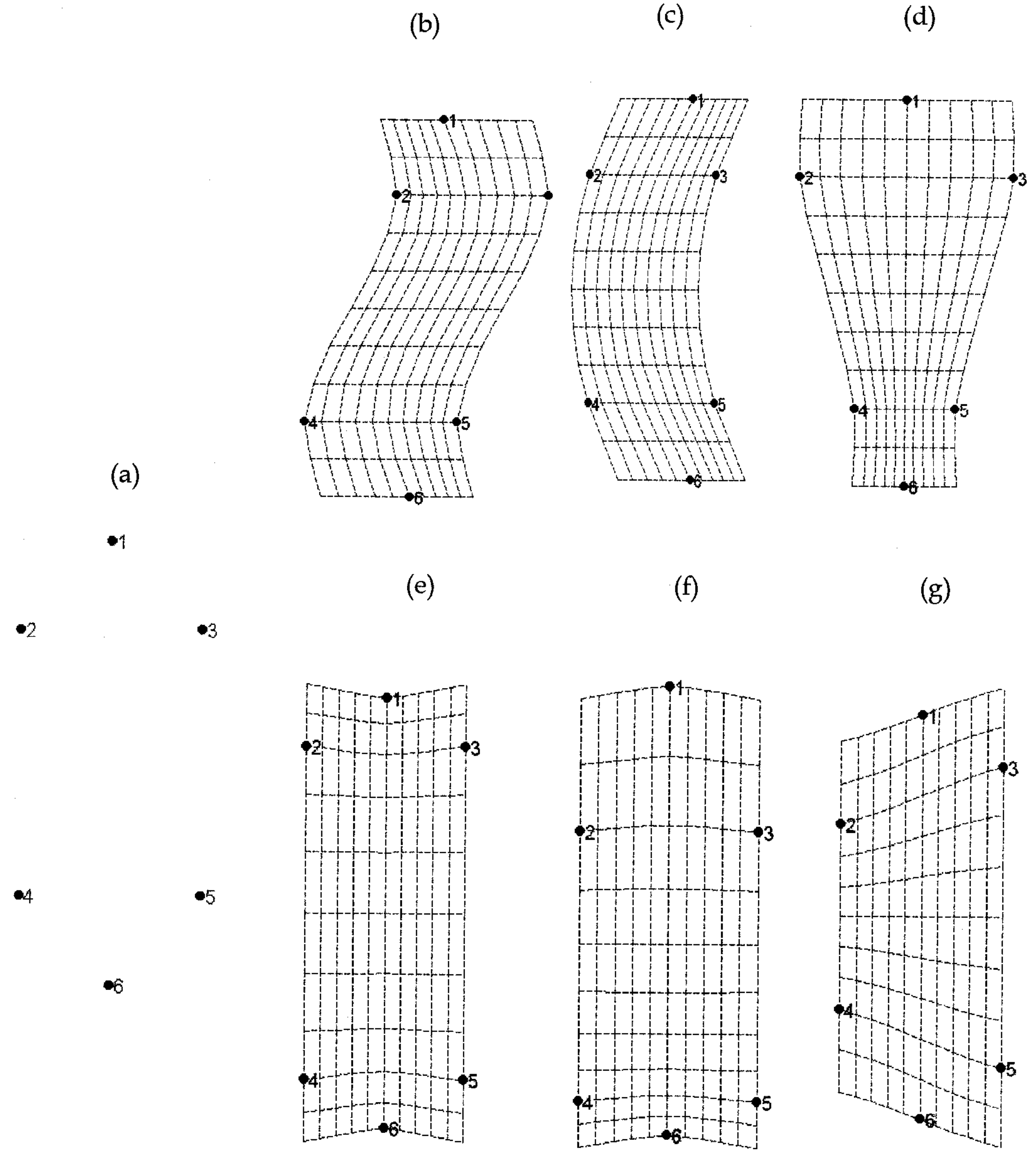

FIGURE 2. A reference configuration (a) and its principal warps (b–g). The warps are ordered from smallest (toward the left) to largest (right) spatial scale and expressed as deformation in x coordinates (b–d) and in y coordinates (e–g). The bending energies are 0.255, 0.140, and 0.125.

fine a space that captures the variation in all possible nonaffine shape variables. Paradoxically, it does this well even though each variable by itself is somewhat arbitrary and need not be especially interesting biologically.

UNIVARIATE VERSUS MULTIVARIATE ANALYSES OF SHAPE

Another important problem with the methods used by Z&F is that they perform statistical analyses on the partial warps one at a time. Although the partial warps

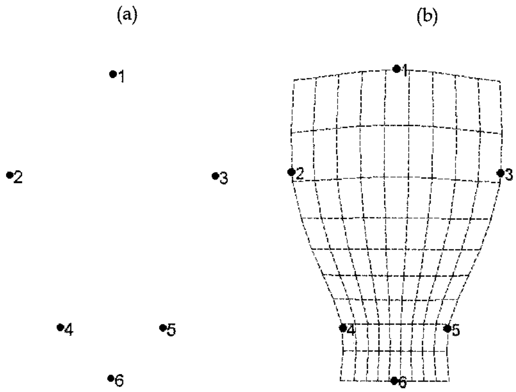

FIGURE 3. (a) Data confirguration in which the size of the region delimited by landmarks 1–3 has been expanded. (b) Thin-plate spline from the reference in Figure 2a to the data configuration.

are geometrically orthogonal, they are not statistically orthogonal—that is, they are not statistically independent. In fact, they tend to be highly correlated. For the rat growth data mentioned earlier, almost half of the correlations between pairs of partial warps were greater than 0.5 and many were larger than 0.8. Performing a series of univariate tests assumes either that the variables are statistically independent or that each variable is of particular biological interest a priori. Partial warps meet neither of these conditions. Standard multivariate statistical methods can be used to take these correlations into account and to make use of the tangent space as a whole.

It is essential that methods be used whose results do not depend upon the particular orientation of the partial warp axes in the tangent space. Thus Bookstein (1996b:146) warns that one should not "believe any multivariate analysis unless its substantive import is independent of the choice of this basis [i.e., the choice of shape variables] . . . Any finding that requires the use of partial warps is erroneous." This means that studies that analyze partial warps must use statistical methods whose results are invariant to the effects of orthonormal transformations of the variables. An orthonormal transformation is a linear transformation of a multivariate space that can be visualized as a rigid rotation of the coordinate axes. Distances between pairs of points and angles between vectors are not effected by such transformations (Reyment and Jöreskog, 1993:46). Note that we are referring here to transformations of the multivariate space of shape variables and not to rotations of the organism. The principal warps are invariant to changes in the orientation of the or-

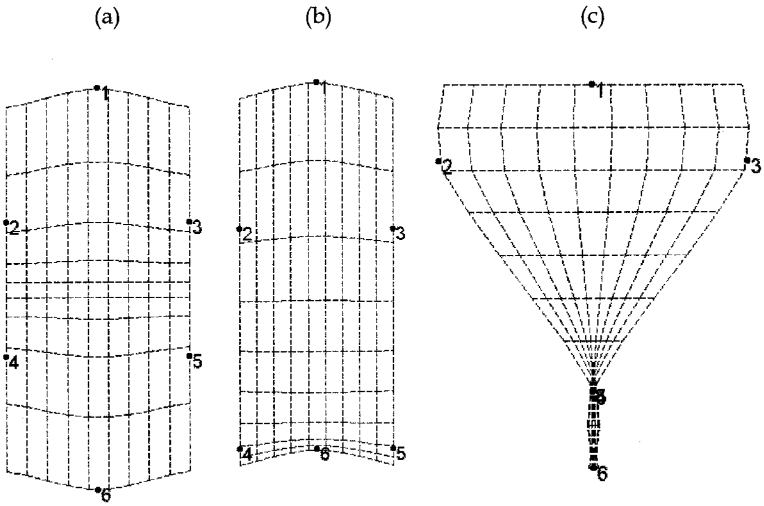

FIGURE 4. Partial warps for the specimen in Figure 3a using the principal warps shown in Figure 2. Despite the fact that the only change was the enlargement of a single region, three partial warps were required to express the change.

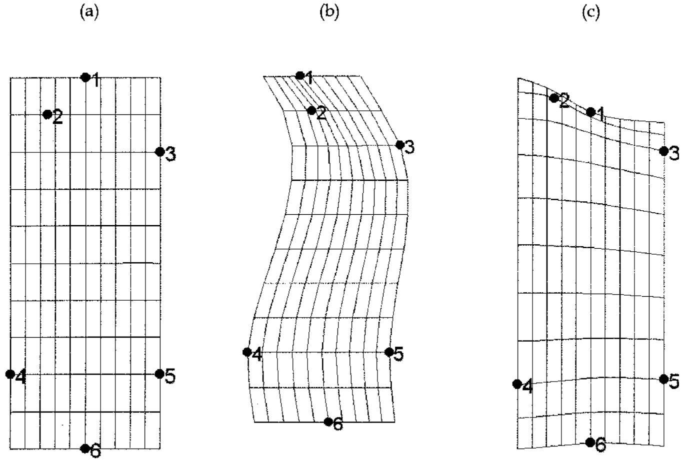

FIGURE 5. The effect of a small change in the reference on the smallest scale principal warp. (a) A reference as in Figure 2a but with landmark 2 moved closer to landmark 1. (b, c) Smallest scale principal warp expressed as a deformation in the x and y coordiinates, respectively. The corresponding shape variable is now a contrast in the relative positions of landmarks 1 and 2, whereas in Figures 2b and 2e it was a contrast in the positions of landmarks 1,4,5 versus 2,3,6.

ganism but the partial warps are not. Rotations of the axes generate new shape variables. Although individually these may or may not be more interpretable than those corresponding to some other rotation, taken as a whole they all capture the same information about shape variation. Thus appropriate statistical methods should lead to the same results. This allows one to use the partial warps as shape variables without assuming them to be especially meaningful one at a time.

The protocols used by Z&F do not share the invariance properties just described. The most important problem is that they examined each principal warp individually to determine its usefulness for yielding partial warps as taxonomic characters. Ontogenetic changes were estimated for each warp by performing multiple regression analyses to predict size as a function of the pairs of x and y components of each partial warp. Although an overall test for ontogenetic shape change can be made by regressing size onto all of the partial warps, separate regression of size onto pairs of partial warps assumes that each principal warp corresponds to a biologically meaningful unit. When significant relationships were found, they regressed each partial warp onto size. The results of these regression analyses were reduced to discrete states so they could be analyzed by parsimony analyses. This was done as a function of the existence of statistically significant regression coefficients and their signs. The effects of breaking up continuous variables (differences between regression slopes) into discrete states is also not invariant to rotations of axes.

On the other hand, the results of analyses such as principal components analysis (PCA) and clustering of Euclidean distance matrices are invariant to rigid rotations of axes. Methods such as multivariate analysis of variance (MANOVA), Hotelling's generalized T^2 tests, canonical variates analysis (CVA), multivariate regression, and Mahalanobis distances are invariant not only to rigid rotations but also to affine transformations as well (they are, for example, also invariant to the effects of multiplication of the axes by scale factors). Rohlf et al. (1996) give a number of examples of the application of these methods to geometric morphometrics.

Multivariate analyses can also be used to construct empirically interesting shape variables. For example, one can construct statistically orthogonal sets of variables that account for the majority of the variance in a few dimensions (principal components) or that best distinguish between two groups (discriminant functions). There are a number of ways to display such shape variables. A visually appealing method is to construct hypothetical shapes using the thin-plate spline. Rohlf et al. (1996) show several examples. Computer programs are available from the Stony Brook Morphometrics www pages (http://life.bio.sunysb.edu/morph) that provide these visualizations for a variety of statistical analyses.

On the other hand, one might argue that in cladistic studies one is not interested in distances but only in character states. In that case, it seems illogical to base the characters on partial warps because partial warps are based on a geometry that depends on Procrustes distances between pairs of shapes.

DISCUSSION

Bookstein (1994) pointed out a number of theoretical problems in applying morphometric methods in character-based cladistic analyses. These problems are consequences of the fact that shape space is curved and high dimensional. Although his arguments are somewhat abstract, the problems are real and cannot be ignored. Two problems are of particular interest and are discussed here.

First, due to the fact that Kendall's (1981, 1984) shape space is curved, the meaning of a shape change resulting from a specified change in the value of a shape variable differs depending upon the values of other shape variables. For example, a change in shape variable *A* from a value of 1 to 2 followed by a change in variable *B* from a value of 0 to 1 does not yield the exact same final shape that would result if the change was in variable *B* followed by the change in variable *A*. Thus the biological meaning of a shape change in a shape variable depends upon the values of the other shape variables. Bookstein (1994) discussed this lack of commutivity of shape transformations. He also discussed a related problem that he referred to as the nonexistence of rectangles in shape space. These properties of shape variables are incompatible with methods that treat variables and their states as entities that can be treated as biologically meaningful units.

A second problem Bookstein (1994) referred to as the shape nonmonotonicity theorem. There is an infinite number of different arbitrary rotations of shape space. Each rotation leads to shape variables that can have different ordering and spacing of points along particular axes. For three noncollinear points in two-dimensional space (2D) or four noncollinear points in 3D one can easily rotate the space to find a shape variable that gives any ordering of the OTUs that one wishes. For more landmarks there are geometric constraints (e.g., interior landmarks cannot be at either end of the orderings), but a great many orderings are still possible. Thus, as discussed earlier, methods need to take into account the facts that partial warps need not be of particular biological interest and that variables generated by different rotations of the tangent space may be of equal or greater interest. Although Z&F were successful in finding taxonomically useful characters, that does not mean that the partial warps they used are especially meaningful. Other linear combinations of the partial warps (e.g., rotations resulting from slightly different choices of the reference configuration) are just as likely to be interpretable. Of course, one may find taxonomically useful characters serendipitously. All directions in tangent space correspond to shape variables, and it is likely that most can be interpreted in some way. In order to use shape variables with a protocol that is not invariant to rotation one needs to justify that the selected variables are each of particular biological interest in comparison to an infinite number of other possible shape variables. Variables based on empirical patterns of covariation in the data are more likely to be biologically meaningful than those generated by an arbitrary a priori rotation of shape space.

An analogy to the use of Fourier analysis (Lestrel, 1997) to study variation in outline shape may be helpful for those already familiar with that approach. Fourier coefficients give the weights for the contribution of the sine and cosine terms for each harmonic just as partial warp scores give the weights for each principal warp. Given a sufficient number of harmonics and carefully aligned outlines, Fourier coefficients can be used as coordinate axes for a space in which points correspond to shapes of entire outlines (Rohlf, 1990, 1993a) just as partial warp scores can be used as coordinate axes for a shape space where points correspond to entire landmark configurations. The different harmonics refer to different spatial scales, as do the principal warps. One should not try to give biological interpretations to the individual harmonics (Bookstein et al., 1982; Rohlf, 1993a), because they are a priori defined variables that do not take patterns of covariation into account. The individual partial warps should not be interpreted for similar reasons. The relationship between Fourier coefficients and the physics of vibrating strings is irrelevant to their use for the analysis of outline shape. Similarly, the relationship between partial warp scores and the physics of deformations of thin metal plates is irrelevant to their use for the analysis of the shapes of configurations of landmarks.

The comments just expressed should not be interpreted as implying that geometric morphometric methods cannot be used in evolutionary biology. Analyses must take into account that partial warps are continuous variables and the orientations of the partial warp axes are arbitrary biologically. Bookstein's (1994) critique is mostly aimed at attempts to use partial warps as characters in cladistic studies and on the use of statistical distance coefficients to measure dissimilarity between shapes. It should also be clear that the present paper is concerned with general mathematical and statistical issues that have nothing to do with the phenetics/cladistics controversies, because we are not concerned here with methods for creating classifications.

For phylogenetic estimation one could use Felsenstein's (1981) maximum-likelihood method for continuous data or Maddison's (1991) squared-change parsimony methods, because they are invariant to the effects of rotation. The maximum-likelihood method is particularly appropriate, because the Procrustes metric used to define shape space is consistent with modeling morphological change as a multidimensional random walk. Methods based on Manhattan distances or those that require the reduction of continuous variables into discrete states are not appropriate for geometric morphometric variables, because their results are not invariant to the effects of arbitrary rotations of shape space. A limitation to the use of shape space to estimate phylogeny is the fact that morphometric analyses typically are limited to a single structure whose components are often highly correlated. Thus shape space often represents just a few independent characters (however defined). It will probably prove to be much more useful to fit an estimated phylogeny obtained from other data and study how shapes change along the estimated lineages (see Bookstein, 1994:211).

The field of geometric morphometrics has reached a level of maturity where problems in ecophenotypy and evolutionary biology can now be investigated. The studies must be based, however, on a firm understanding of the methods being used to minimize the possibility of biases and statistical artifacts in the results. The new methods are very powerful and must be used with care.

ACKNOWLEDGMENTS

The insightful criticisms by Fred L. Bookstein and the helpful comments by Dean Adams, Mike Rosenberg, and Dennis E. Slice are gratefully acknowledged. This work was supported by a grant (DEB 93-17572) from the Systematic Biology Program of the National Science Foundation. This paper is contribution no. 997 from the Graduate Studies in Ecology and Evolution, State University of New York at Stony Brook.

REFERENCES

ADAMS, D. C., AND M. S. ROSENBERG. 1998. Partial warps, phylogeny, and ontogeny: A comment on Fink and Zelditch (1995). Syst. Biol. 46:168–173.

BOOKSTEIN, F. L. 1989. Principal warps: Thin-plate splines and the decomposition of deformations.

IEEE Trans. Pattern Anal. Machine Intelligence 11: 567–585.

BOOKSTEIN, F. L. 1991. Morphometric tools for landmark data. Cambridge Univ. Press, New York.

BOOKSTEIN, F. L. 1994. Can biometrical shape be a homologous character? Pages 197–227 *in* Homology: The hierarchical basis of comparative biology (B. K. Hall, ed.). Academic Press, New York.

BOOKSTEIN, F. L. 1996a. Biometrics, biomathematics and the morphometric synthesis. Bull. Math. Biol. 58:313–365.

BOOKSTEIN, F. L. 1996b. Combining the tools of geometric morphometrics. Pages 131–151 *in* Advances in morphometrics (L. F. Marcus, M. Corti, A. Loy, G. Naylor, and D. E. Slice, eds.). Plenum, New York.

BOOKSTEIN, F. L. 1996c. Standard formula for the uniform shape component in landmark data. Pages 153–168 *in* Advances in morphometrics (L. F. Marcus, M. Corti, A. Loy, G. Naylor, and D. E. Slice, eds.). Plenum, New York.

BOOKSTEIN, F. L., B. CHERNOFF, R. ELDER, J. HUMPHRIES, G. SMITH, AND R. STRAUSS. 1982. A comment on the uses of Fourier analysis in systematics. Syst. Zool. 31:85–92.

BURKE, R. L., T. E. LEUTERITZ, AND A. J. WOLF. 1996. Phylogenetic relationships of emydine turtles. Herpetologia 52:572–584.

FELSENSTEIN, J. 1981. Evolutionary trees from gene frequencies and quantitative characters: Finding maximum-likelihood estimates. Evolution 35:1229–1242

FINK, W. L., AND M. L. ZELDITCH. 1995. Phylogenetic analysis of ontogenetic shape transformations: A reassessment of the piranha genus *Pygocentrus* (Teleostei). Syst. Biol. 44:344–361.

KENDALL, D. G. 1981. The statistics of shape. Pages 75–80 *in* Interpreting multivariate data (V. Barnett, ed.). Wiley, New York.

KENDALL, D. G. 1984. Shape-manifolds, Procrustean metrics and complex projective spaces. Bull. Lond. Math. Soc. 16:81–121.

LESTREL, P. E. 1997. Fourier descriptors and their applications in biology. Cambridge, New York.

LYNCH, J. M., C. G. WOOD, AND S. A. LUBOGA. 1996. Geometric morphometrics in primatology: Craniofacial variation *Homo sapiens* and *Pan troglodytes.* Folia Primatol. 67:15–39.

MADDISON, W. P. 1991. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. Syst. Zool. 40: 304–314.

NAYLOR, G. J. P. 1996. Can partial warp scores be used as cladistic characters? Pages 519–530 *in* Advances in morphometrics (L. F. Marcus, M. Corti, A. Loy, G. Naylor, and D. E. Slice, eds.). Plenum, New York.

REYMENT, R. A., AND K. G. JÖRESKOG. 1993. Applied factor analysis in the natural sciences. Cambridge, New York.

ROHLF, F. J. 1990. Fitting curves to outlines. Pages 167–177 *in* Proceedings of the. Michigan morphometrics workshop (F. J. Rohlf and F. L. Bookstein, eds.). Univ. Michigan Museum of Zoology (Special Publication no. 2), Ann Arbor.

ROHLF, F. J. 1993a. Feature extraction applied to systematics. Pages 375–392 *in* Advances in computer methods for systematic biology (R. Fortuner, ed.). Johns Hopkins Univ. Press, Baltimore.

ROHLF, F. J. 1993b. Relative warp analysis and an example of its application to mosquito wings. Pages 131–159 *in* Contributions to morphometrics (L. F. Marcus, E. Bello, A. Garcia-Valdecasas, eds.). Museo Nacional de Ciencias Naturales (CSIC), Vol. 8. Madrid, Spain.

ROHLF, F. J. 1996. Morphometric spaces, shape components, and the effects of linear transformations. Pages 117–129 *in* Advances in Morphometrics (L. F. Marcus, M. Corti, A. Loy, G. Naylor, and D. E. Slice, eds.). Proceedings of the 1993 NATO Advanced Studies Institute on Morphometrics in Il Ciocco, Italy. Plenum, New York.

ROHLF, F. J. 1997. TpsSmall. Department of Ecology and Evolution, State Univ. New York, Stony Brook, NY.

ROHLF, F. J., LOY, A., AND M. CORTI. 1996. Morphometric analysis of Old World Talpidae (Mammalia, Insectivora) using partial warp scores. Syst. Biol. 45: 344–362.

ROHLF, F. J., AND L. F. MARCUS. 1993. A revolution in morphometrics. Trends Ecol. Evol. 8:129–132.

ROHLF, F. J., AND D. SLICE. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst. Zool. 39:40–59.

SWIDERSKI, D. L. 1993. Morphological evolution of the scapula in tree squirrels, chipmunks and ground squirrels (Sciuridae): An analysis using thin-plate splines. Evolution 47:1854–1873.

ZELDITCH, M. L., F. L. BOOKSTEIN, AND B. L. LUNDRIGAN. 1992. Ontogeny of integrated skull growth in the cotton rat *Sigmodon fulviventer.* Evolution 46: 1164–1180

ZELDITCH, M. L., F. L. BOOKSTEIN, AND B. L. LUNDRIGAN. 1993. The ontogenetic complexity of developmental constraints. J. Evol. Biol. 6:121–141.

ZELDITCH, M. L., AND W. L. FINK. 1995. Allometry and developmental integration of body growth in a piranha, *Pygocentrus nattereri* (Teleostei: Ostariophysi). J. Morphol. 223:341–355.

ZELDITCH, M. L., W. L. FINK, AND D. L. SWIDERSKI. 1995. Morphometrics, homology, and phylogenetics: Quantified characters as synapomorphies. Syst. Biol. 44:179–189.

Received 18 March 1997; accepted 14 June 1997
Associate Editor: D. Cannatella