coordinate data and standard multivariate statistical analyses which are available in most major statistical packages. At the University of Washington we have found it especially convenient to carry out our analyses using an interactive statistical programming language such as "S" (Becker and Chambers, 1984), or "ISP" (Dunlap, 1985), both of which have facilities for user-defined special purpose macros. In this interactive macro environment we easily extract shape coordinates (using the simple expressions of complex arithmetic given by Bookstein) for arbitrary sets of landmarks, compute the usual statistical analyses, and generate various graphical displays of the results.

One of Bookstein's most important contributions to the field of morphometrics was the method of biorthogonal grids which he introduced in 1978. We are finding biorthogonal grids very useful for graphically synthesizing the findings from the discrete analyses of multiple triangles (as Bookstein describes in Section 6). However, to our knowledge no one but Bookstein himself at the University of Michigan has ever had software to generate a biorthogonal grid. This is probably due to the complexity of the algorithms originally described. We have recently implemented (with Bookstein), in the "S" environment, new and simpler algorithms for the computation of biorthogonal grids. The computed homology which maps and smoothly interpolates one set of landmarks onto another is derived from easily programmed "thin-plate" spline interpolators (Meinguet, 1979). This algorithm does not constrain the mapping to be linear on a specified boundary as does Bookstein's

original algorithm. Our algorithm for drawing out the biorthogonal grids, the integral curves of the symmetric tensor field (Figure 15b), is based on a widely available differential equation solver. A report describing this new biorthogonal grid software and applications is in preparation.

ACKNOWLEDGMENTS

ADDITIONAL REFERENCES

BECKER, R. A. and CHAMBERS, J. M. (1984). *S: An Interactive Environment for Data Analysis and Graphics.* Wadsworth, Belmont, Calif.

DUNLAP, W. W. (1985). ISP User's Guide. Associated Students of the Univ. of Washington Lecture Notes, Seattle.

MEINGUET, J. (1979). Multivariate interpolation at arbitrary points made simple. *J. Appl. Math. Phys.* **30** 292–304.

SAMPSON, P. D. and SIEGEL, A. F. (1984). The measure of 'size' independent of 'shape' for multivariate lognormal populations: definition and applications. *Proceedings of the XIIth International Biometric Conference, Tokyo.*

SAMPSON, P. D. and SIEGEL, A. F. (1985). The measure of 'size' independent of 'shape' for multivariate lognormal populations. *J. Amer. Statist. Assoc.* **80** 910–914.

STREISSGUTH, A. P., MARTIN, D. C., MARTIN, J. C. and BARR, H. M. (1981). The Seattle longitudinal prospective study on alcohol and pregnancy. *Neurobehav. Toxicol. Teratol.* **3** 223–233.

Comment

Colin Goodall

Fred Bookstein's energy, enthusiasm, leadership, and innovative thinking about morphometrics are highly valued, greatly appreciated, and a spur to further work. The present paper is a major advance in multivariate morphometrics, and contains some of the few substantive results available. The linear spaces for size and shape statistics are derived, however, at the cost of restrictive assumptions, including a simple error structure (the null model), almost uniform deformation (negligible curvature), and small errors

Colin Goodall is Assistant Professor, Engineering Statistics and Management Science Department, Princeton University, E-Quad, Princeton, New Jersey 08544.

(linearized, normally distributed, statistics). This discussion looks at a broader approach, and, while lacking the detail and rigor of Bookstein's paper, suggests that statistical machinery, centered on function estimation, is mostly available.

The author has convincingly demonstrated how to move back and forth between deformations and multivariate statistics. These statistics are based directly on linear combinations of landmarks. I prefer to emphasize a two-stage procedure, in which estimation of the biological process, namely the deformation (strain) tensor field varying in space and time, is primary. Only at the second stage statistics that summarize (are functionals of) the deformation tensor field are used in multivariate comparisons. As Bookstein

discusses, this approach follows Thompson (1942, 1961), Sneath (1967), Bookstein (1978a), and others.

The first stage involves fitting of the deformation tensor field to the sets of landmark co-ordinates. This bears a close resemblance to fitting curves (or lines) through bivariate (x, y) data, but is more complex. For two-dimensional deformations there are two x co-ordinates (before deformation), two y co-ordinates, and interest focuses on the 2×2 derivative (tensor, Jacobian) of the mapping from **x** to **y** As with curve fitting to bivariate data there are a number of alternatives: parametric and nonparametric, interpolation and smoothing, regression and errors-in-variables. A parametric mapping may be a similarity transformation (Siegel and Benson, 1982), a linear (affine), quadratic, or cubic transformation. A single parametric transformation may be fit to all landmarks, or separate transformations to each of several subsets of landmarks. The subsets may correspond to a cellulation (piece-wise linear fit say, with as special case a triangulation), or overlap arbitrarily (Bookstein considers *all* subsets of landmarks). For a review, see Goodall (1984).

Statistics for multivariate analyses are most easily obtained from parametric fits. As is the case generally with data smoothing, parametric fits are efficient when the bias is kept small. For initial, exploratory analysis of the deformation of an organism a nonparametric approach is recommended. As Bookstein points out in his motivation of the factor analysis paradigm, the deformation tensor field is real. It is also nonconstant in space (and time).

A promising notion, as Bookstein has remarked in conversation, is shape invariant modelling (SIM, Stuetzle et al., 1980). SIM avoids some problems by allowing a nonparametric fit related across individuals by parameters that will usefully (see below) be in terms of geometrical invariants, the principal strains and directions.

To help further understand the alternatives, three sources of variation may be distinguished.

1. Measurement error of landmark locations, before and after deformation.

2. Modelling error or bias, specializing to inhomogeneity error when the model is an homogeneous affine transformation.

3. Interindividual variation.

Some qualitative distinctions must be made between the three sources of variation. Plausibly, measurement error is independent at each landmark. But, as Bookstein points out, landmark coordinates (the Z_i) are defined relatively, and not absolutely. The Cartesian coordinate system has relevance as soon as the organism is positioned on a (perhaps metaphorical) piece of graph paper for digitizing. The transformation of interest is defined to within a rigid body motion.

However, inhomogeneity error and interindividual variation exist prior to this positioning, and should be modelled directly in the growth process, i.e., in the deformation tensor field. The three sources of error are now discussed in more detail.

Measurement error may be circular normal, or contaminated with outliers, short-tailed, or nonisotropic reflecting pixel shape and irregularities in the image near the landmark (Goodall, 1984). Measurement error, found in landmarks both before and after deformation, demands an errors-in-variables approach, as in Gleser and Watson (1973) for three-dimensional data (Bookstein presents a highly linearized version).

We may model inhomogeneity error as random errors in landmark coordinates *after* deformation only, suggesting a regression approach when measurement error is relatively small. An example is a single affine transformation fit to the landmarks. Let $\mathbf{x}_i$ and $\mathbf{y}_i$ be homologous landmark coordinates before and after deformation, where $i = 1, \cdots, n$, $n \geq 3$ landmarks. Let A be a 2×2 affine transformation matrix, and $\mathbf{e}_i$ a residual vector. With a variety of least-squares fitting procedures (regression, errors in variables, canonical correlations), the centroid of the landmarks before deformation, $\bar{\mathbf{x}}$, transforms to the centroid after deformation, $\bar{\mathbf{y}}$. We write

$$(1) \qquad \mathbf{y}_i = A\mathbf{x}_i + (\bar{\mathbf{y}} - A\bar{\mathbf{x}}) + \mathbf{e}_i$$

for the bidimensional regression model. A more realistic approach may involve weights, decreasing with distance from the landmark centroid, and spatially autocorrelated errors.

More specialized still is the exact fit of the (six parameter) homogeneous linear (affine) transformation to three landmarks. The exactness is misleading: zero residuals do not imply zero errors or an homogenous tensor field. There are no residual degrees of freedom, and no estimate of the measurement or inhomogeneity error variance from a *single* triangle of landmarks. An analog for bivariate data is to take slopes of line segments joining pairs of points. With its constant reference to triangles, Bookstein's approach does not model inhomogeneity error. He explicitly assumes that curvature is negligible.

Interindividual variation is properly modelled in the geometrical invariants (the principal strains), and their directions, of the symmetric deformation tensor field. Consider the 2×2 affine transformation matrix, A. By a singular value decomposition, S decomposes into a rotation, dilation and rotation,

$$(2) \qquad A = R_\psi D_{p,q} R_{-\theta},$$

where R_ψ is rotation through an angle ψ and $D_{p,q} = diag(p, q)$. The principal strains are p and q, and the maximum strain (p) is at an angle θ in the coordinate frame before deformation, at angle ψ after

deformation. The symmetric tensor is

$$R_\theta R_{p,q} R_{-\theta}.$$

(An intermediate log transformation of p and q gives strain rates.)

Interindividual variation is modelled directly in p, q (or $\ln p$, $\ln q$), θ, and ψ, as normal perturbations about a population mean. From equations (1) and (2), $\mathbf{y}_i$ is a linear function of p and q. The variance of $\mathbf{y}_i$ varies with the projection of $\mathbf{x}_i - \bar{\mathbf{x}}$ on the principal axes of deformation. The deformation is shared by landmarks, introducing correlations. The covariance of $\mathbf{y}_i$ and $\mathbf{y}_j$ includes the expression $(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})^T$. Invoking linearization, the $\mathbf{y}_i$ are also approximately normally distributed when a) $\ln p$ and $\ln q$ and b) θ and ψ are normal with small variance. Normal errors in the principal strains and directions are propagated (approximately) linearly, giving a setup differing from Bookstein's null model because heteroscedastic and correlation effects are a necessary structural feature.

Although the principal strain invariants are properly comparable across individuals at each landmark, modelling the distributions of θ and ψ requires a benchmark direction at each point, common across individuals. A further difficulty is to make comparisons across individuals at points other than landmarks, for which pseudo-landmarks are needed. These issues are considered again at the end of the discussion.

Interindividual variation may be grouped with inhomogeneity error in a regression-like model. These two sources of error differ in that interindividual variation has a well defined covariance structure (via equation (1)). A variance components approach is needed to separate out the three sources of error, but there is still an identifiability problem. Interindividual variation is modelled on the true deformation tensor field for each organism, and therefore inhomogeneity error is confounded with it. Systematic residuals across individuals indicates inhomogeneity error, but interindividual variation can reverse the sign of any modelling bias.

With regard to assumptions, the observed normality of the $d\mathbf{Q}$ is reassurance that an a priori normal model is justifiable. It does not assert or prove normality when there is no a priori argument for normality.

As a general remark, size (S^2) and shape statistics are uncorrelated in an errors-in-variables, measurement-error approach. It has not been worked out how errors in the deformation parameters affect this.

A goal at the second stage is to make available a large class of summary statistics of the deformation. As a general rule, these statistics should be geometrically meaningful, so will usually involve the eigenvalues and eigendirections of the strain tensor, p, q, θ, and ψ, rather than parameters of the fit, and will

measure change of size and shape. This may be at a specific landmark, the extrema over the form, or involve integration over an area (for example, a triangle of landmarks) or along a path (the straight line between two landmarks, or an integral curve of the biorthogonal grid). More ambitious are confidence bounds for the tensor field and functionals of it. I am less pessimistic than the author appears to be about technology for statistical comparisons based on the deformation. Any deformation can be decomposed to give the principal strains and directions at each point.

The size statistics discussed in the paper are a particular subclass of the above, namely those belonging to the linear subspace generated by the integration between pairs of landmarks of an *interpolated* deformation tensor field. The linear space of shape statistics has as basis the ratios of two basis size statistics sharing a landmark, so should involve integration over triangles of landmarks of an interpolated deformation tensor field. There are a number of difficulties:

1. In general there does not exist an interpolated deformation tensor field such that the ratio of distances between any pair of pseudo-landmarks is the appropriate integral. The problem is that the same coefficients c_i are used in a contrast before and after deformation, i.e., each pseudo-landmark is a true landmark, and this assumes an everywhere linear deformation.

2. Smoothing is more appropriate than interpolation. While admitting the existence of measurement error, Bookstein's approach does nothing to reduce its impact and separate it from interindividual variation. An intelligent smoothing algorithm will incorporate the variance components considerations alluded to above, and reduce measurement error while retaining interindividual variation.

3. The author's Theorem 2, pertaining to the existence of the extremum admissible size variable in the set of size variables generated by triangles of landmarks alone, is a powerful one. Reservations, however, stem from the assumption of zero curvature within each triangular area. Extremum size change is the extremum over all triangles of landmarks but may differ (due to curvature) from the pointwise extremum. Similarly, the linearized space of shape variables depends on the affine differentials for triangles of landmarks. The basis depends critically on the landmark configuration, which may be sparse in the regions of greatest interest. Higher order interpolation and smoothing will provide information on local extrema.

It is hard to choose the right functionals of the deformation tensor field, but it is also important to allow for curvature and the three sources of error. Any choice of functionals, to be tractable, involves loss of information. The Bookstein alternative is attractive

if we believe the proposed null model, expect small changes and negligible curvature, and emphasize landmarks. It is useful and valuable, but does not go far enough.

I contend that multivariate statistical methods may be based directly on functionals of the deformation tensor field calculated for each organism. The three sources of variability inter-relate in the eventual analysis. One extreme case is normal inhomogeneity error and negligible measurement error. Then the 2×2 transformation matrix of a single affine transformation fit to more than 3 landmarks using a bidimensional regression technique (Tobler, 1978) has a bivariate normal distribution (Goodall, 1984). Tests for isotropy and directionality follow immediately. Within and between group comparisons require additional assumptions about the population of true affine transformations. Another extreme case is when measurement and inhomogeneity error are both negligible. As discussed above, the appropriate approach is governed directly by our assumptions about the distribution of the deformation tensor field across individuals.

An incidental benefit of the two-stage approach is that the factor analysis paradigm is less central. Factor analysis, while it does provide an elegant analog in accounting for the covariation of distance measures, is nevertheless a technique to be viewed with caution. Fortunately, for morphometric purposes many of the estimation problems are avoided because the factors are real and of known dimensionality, namely they are the deformation tensor field.

There is a further intriguing aspect. Bookstein's null model assumes that normal errors in landmark coordinates include interindividual variation. The result is an approximate normal distribution for the population factor scores, i.e., he implicitly assumes a normal population model above the factor analysis one.

The shape statistic of Bookstein, $d\mathbf{Q}$, has an algebraic interpretation related to the affine transformation matrix and its singular value decomposition given in equations (1) and (2) above. Of particular interest is the approximate result that $|d\mathbf{Q}| \approx (\ln p - \ln q)h$, where h is the height of the triangle.

Suppose that the edge $\mathbf{x}_1\mathbf{x}_2$ has length $\alpha = \| \mathbf{x}_2 - \mathbf{x}_1 \|$ and orientation α. Then the similarity transformation

$$(3) \qquad \mathbf{x}' = \frac{1}{\alpha} R_{-\alpha}(\mathbf{x} - \mathbf{x}_1)$$

standardizes $\mathbf{x}_1$ and $\mathbf{x}_2$ to $(0, 0)$ and $(1, 0)$. With b and β defined analogously,

$$(4) \qquad \mathbf{y}' = \frac{1}{b} R_{-\beta}(\mathbf{y} - \mathbf{y}_1)$$

standardizes $\mathbf{y}_1$ and $\mathbf{y}_2$ to $(0, 0)$ and $(1, 0)$.

For $n = 3$, $\mathbf{e}_i \equiv \mathbf{0}$ and $\mathbf{y}_1 - A\mathbf{x}_1 = \bar{\mathbf{m}} - A\bar{\mathbf{x}}$. Therefore the affine ratio $\mathbf{Q}$ is equivalent to the affine transformation

$$(5) \qquad \mathbf{y}' = A'\mathbf{x}'$$

where

$$(6) \qquad A' = R_{\psi-\beta} D_{pa/b, qa/b} R_{-(\theta-\alpha)}.$$

Then

$$(7) \qquad d\mathbf{Q} = (A' - I)\begin{pmatrix} r_1 \\ s_1 \end{pmatrix}$$

$$(8) \qquad = \begin{pmatrix} \frac{a^2}{b^2}(p^2 - q^2)\frac{1}{2} Sh \\ \left(\frac{a^2}{b^2} pq - 1\right)h \end{pmatrix}$$

where r_1 and s_1 are as defined in the paper, $S = \sin 2(\theta - \alpha)$, and $h = \operatorname{Im} \mathbf{Q} = s_1$. By the familiar argument,

$$(9) \qquad \frac{b^2}{a^2} = p^2\cos^2(\theta - \alpha) + q^2\sin^2(\theta - \alpha)$$

and

$$(10) \qquad d\mathbf{Q} = \begin{pmatrix} \dfrac{(p^2 - q^2)Sh}{(p^2 + q^2) + C(p^2 - q^2)} \\ \dfrac{-(p - q)^2 - C(p^2 - q^2)}{(p^2 + q^2) + C(p^2 - q^2)}h \end{pmatrix}$$

where $C = \cos 2(\theta - \alpha)$.

Let $q = p(1 - \delta)$, where $\delta \ll 1$. Then

$$(11) \qquad d\mathbf{Q} \approx \begin{pmatrix} S \\ -C \end{pmatrix}\delta h.$$

Then $d\mathbf{Q}$ has length δh and direction $2(\theta - \alpha) - \pi/2$. These results confirm those of Bookstein (1984a) cited in the text above. (The additional log transformation does not affect the linearized results.) A number of comments follow.

1. For small strains permutation of landmarks in the triangle rescales $d\mathbf{Q}$ by the ratio of h's, and rotates $d\mathbf{Q}$ by twice the angle between the edges. Thus population differences are disturbed by the stochastic variation in h and in the angles. A dimensionality argument shows that the difference between (log) principal strains must be weighted by h to obtain $d\mathbf{Q}$. However, h is the height of the triangle *after* standardization by (3), so an "unstandardized $d\mathbf{Q}$" is related to δ by area.

In terms of the model for interindividual variation introduced above, in which the distributions of p and q are specified directly, the same phenomenon arises

in the linear $\mathbf{x}_i - \bar{\mathbf{x}}$ term relating A to $\mathbf{y}_i - \bar{\mathbf{y}}$ (1). With this model the δ should be compared directly, with triangle dimension entering only in estimating the precision.

2. For large strains $d\mathbf{Q}$ is not invariant (to within a rotation) to choice of base edge (10).

3. One difficulty with the approach via triangles is that measurement error is propagated through the choice of a common baseline (in a quadrilateral say) whereas variance considerations suggest averaging. Only for each triangle of landmarks is the translation of a single landmark always equal to the translation of the landmark centroid, $\mathbf{y}_i - A\bar{\mathbf{x}}_1 = \bar{\mathbf{y}} - A\bar{\mathbf{x}}$.

The modelling inconsistencies in using triangles have already been considered. However, this approach does define a basis direction comparable across individuals and does extend the homology between individuals from landmarks to the whole form (using pseudo-landmarks). As mentioned previously, modelling of interindividual variation requires at least the definition of a basis direction, and some tacit notion of extended homology also. These are the principal conceptual impediments to development of a theory of morphometrics along the lines of this discussion. They are no novelty to morphometrics, and it is sub-

stantially to Bookstein's credit that he has dealt with them.

As a final remark, the next generalization is to longitudinal data, for which the positions of a set of landmarks, possibly evolving in time, are recorded at several time points. The deformation tensor field is varying in space and time. Technically, many of the issues are the same, as in fact approaches for the analysis of finite deformation have been borrowed from the analysis of longitudinal data itself.

ADDITIONAL REFERENCES

GLESER, L. J. and WATSON, G. S. (1973). Estimation of a linear transformation. *Biometrika* **60** 525–534.

GOODALL, C. R. (1984). The growth of a two-dimensional figure: strain crosses and confidence regions. *Proc. Statist. Comput. Section, American Statistical Association, Philadelphia, August 1984.*

STUETZLE, W., GASSER, T., MOLINARI, L., LARGO, R. H., PRADER, A. and HUBER, P. J. (1980). Shape invariant modelling of human growth. *Ann. Human Biol.* **7** 507–528.

Rejoinder

Fred L. Bookstein

Five able discussants have persuaded me that my essay, however long already, spent too little space reviewing themes other than its own. Each discussion points out connections between morphometrics and diverse topics both within biometrics and without.

Kendall's shape space Σ_2^3. David Kendall surmises, correctly, that I had not previously encountered his work. Indeed we have approached nearly the same problem from two very different directions. Although permutations and reflections of landmark configurations are prohibited on biological grounds, the algebra of my shape space is still that of a tangent space at the point of his Σ_2^3 corresponding to the mean shape. In the large, the plane of shape coordinates Q represents all of his shape space, except for one point $Z_1 = Z_2$, six times over.

The tensors supply a canonical geometric description of directions in any tangent plane of this space. Also, they lead to a metric geometry throughout the space, with infinitesimal element of distance equal to $|dQ|/\text{Im } Q$, the difference of the log principal strains. The geodesic arcs of this geometry are curves corresponding to triangular shapes whose transformations

from a fixed starting triangle have the same principal axes—the shapes that can be reached by fractional powers of the same affine transformation. The geodesics, then, must be the circles involved in the construction of the principal axes (Figure 11), the circles orthogonal to the real axis. In this metric construction for shape space we recognize one of the classic models of hyperbolic geometry, the Poincaré half-plane (cf. Coxeter, 1965, Section 14.8).

Such a space has negative Riemannian curvature, whereas Kendall's shape space, under the metric inherited from Euclidean distance, has positive curvature. This and other interesting differences between Kendall's geometry and mine derive from the differences between their fields of principal application. My morphometric shape space has a line of metric singularities all down the real axis. Points (triangles) not on this axis may not be transformed into points upon it or across it by any proper affine transformation. That axis, representing triangles of zero area, is the Absolute of the hyperbolic geometry, the locus infinitely far away. Its exclusion expresses the restriction of the deformation model to transformations of