Morphometrics

Morphometrics may be defined as a more or less interwoven set of largely statistical procedures for analyzing variability in size and shape of organs and organisms. Some of the concepts have been generalized to encompass nonbiological problems. An early morphometric analysis was conducted by Blackith (1959) who applied multivariate statistical methods to the basic carapace morphology of grasshoppers and was able to follow the likelihood of the development of a swarming phase in a population by pinpointing morphological changes heralding a population explosion. It can be used to compare the shapes of animals, for example, or to compare the characteristics of diseased and nondiseased organs. The rise of the science of morphometrics as an organized discipline stems largely from the work of P.C. Mahalanobis and his colleagues at the Indian Statistical Institute in the early 1940s, devoted to the anthropometric survey of the United Provinces of India; there is an important and comprehensive summary by Rao [1] of this survey. The workhorse of the Indian project was the generalized statistical distance of Mahalanobis, which was used to produce topological diagrams with links between all categories, using suites of measurements on the heads of subjects. Later, Rao expanded the methodology to encompass analyses by canonical variates, employed as a means of generalizing the Mahalonobis distance. Craniometrics is concerned with identifying landmarks, defined as distances between diagnostic sites, marked off by the coordinates of points located at these sites. For many purposes, such data are adequate for standard multivariate analyses. However, as soon as it is desired to concentrate on the variation in shape of an organism, a particular problem is encountered, namely that in data consisting of interlandmark distances, size and shape are inextricably confounded. Blackith [2, 3] is responsible for introducing multivariate morphometrics as a unified analytical concept in his ecologically oriented studies on swarming locusts. Burnaby [4] seems to have been the first worker to attempt to find a solution to the problem of size and shape confounding by proposing a transformation that put size into one subspace and shape into another. He called this the *method of growth-free discrimination*. A different approach was, however, generally pursued by workers following an ad hoc procedure that relies on reifying (i.e., interpreting) the latent roots and vectors of a matrix of covariances (usually of the logarithms) of distances. By a process of a posteriori reasoning it is usually suggested that the first latent vector is a measure of allometric variation in size and the second and subsequent latent vectors vaguely define measures of variation in shape. Practical use of this technique appears relatively early on in the agricultural literature (Pearce on fruit trees, for example). The results of multivariate analyses of distances are often presented as *ordinations,* that is, scatter-plots, of the scores pertaining to the specimens on canonical axes. Of mainly historical interest is the imaginative treatise on morphological integration by Olson and Miller [5] in which correlations between distances on bones and skulls were interpreted hierarchically within a morphometric context.

Lohmann's [6] method of "eigenshapes" interposed a more shape-oriented motif into morphometrics; at the time of its inception, it constituted a remarkable achievement. His eigenshapes are actually the first derivative of the outline of the squared differences which, at arbitrarily identified "corresponding" points, are integrated to form a net distance. Another type of outline procedure consists of the application of trigonometric Fourier coefficients to characterize shape contours. So-called Fourier analysis has a similar idea underpinning its application to outline data. Fourier coefficients supply the weights for the contribution of the sine and cosine terms for each harmonic. Given a sufficient number of harmonics, and carefully aligned outlines, Fourier coefficients can be used as coordinate axes for a space in which the points correspond to the shapes of complete outlines. It is not an appropriate procedure to give biological interpretations to the individual harmonics because they are a priori defined variables that do not take account of covariation.

Geometric Morphometrics

Geometric morphometrics is a relatively recent development that provides reasonable solutions to many

Based in part on the article "Morphometrics" by Richard A. Reyment, which appeared in the *Encyclopedia of Environmetrics.*

of the problems and lack of consistency encountered in the past in morphometrical work, for example, the confounding of size and shape. It can be seen as being the mathematical justification of Thompson's [7] coordinate grids and is, in its biologically oriented formulation, largely due to Bookstein [8, 9] and other publications of that author in which exceptional geometric insight has been wedded to practical utility.

The mainstay of geometric morphometrics are data obtained from "landmarks." Landmarks are defined as recognizable equivalent points observed on the objects being compared. Sets of landmarks occur in two or three dimensions; the term "form" may be used to describe the configuration of the set of landmarks. Geometric morphometrics encompasses methods for the analysis of such configurations operating in a specific shape space [10, 11], with full geometry being preserved. The term "shape" refers to the aspects of form remaining with size extracted. Landmark coordinates have the advantage that they register equivalences or homologies in a way that points assigned to outlines do not. Landmarks may be referred to three categories [12]:

Type 1 landmarks have a homology that is supported from case to case by the strongest local evidence (abutting structures or tissues, histological evidence, etc.). This type is the most reliable for analysis

Type 2 landmarks are not supported by local or histological evidence, but are geometric in nature, for example, ornamental features on a carapace, wing-tips. Such landmarks are usually functionally equivalent.

Type 3 landmarks have at least one deficient coordinate and therefore can be located to an outline or a surface but not at a specific location, for example, the tip of a caudal process. This type is the least reliable of the three, but is nonetheless a valuable adjunct in the analysis of difficult material.

There are two main tacks for the localization of differences in form by means of coordinates. One of these takes each form, superimposes it with respect to others, and then records differences in terms of landmark displacements relative to this registration. The second approach describes differences in landmark configurations as deformations of a grid obtained by mapping one form into another and visualizing the compressions and stretchings that are generated by the procedure. The analysis of landmark registrations may be done in several ways, three of which are of main interest. One may register to a common baseline by translating, rotating, and scaling, so that most points fit well, or register by minimizing the sum of squared differences between the equivalent landmarks of forms. This latter procedure is known as *generalized Procrustean fitting* (cf Reyment [13]). Scaling is carried out via centroid size (to wit, the sum of squared Euclidean distances from each landmark to the centroid – the mean of landmark coordinates). Geometric morphometrics has reached a level of progress that now makes it possible to investigate questions of ecophenotypy and polymorphism [14]. It is however essential that analyses be founded on a clear understanding of what the methods invoked actually do and what they cannot do in order to avoid the proliferation of statistical artifacts in the results.

Shape Spaces for Landmarks

Kendall *et al*. ([15], p. 1) espouse a rather rigid interpretation of labeled points (they adroitly bypass the use of the word "landmark") in their mathematical concept of markers in that they underline that for them, labeled points are basic and determine the objects studied. According to the biologically oriented concept of the "geometric morphometricians" the "marker points" are selected from a usually two- or three-dimensional continuum. The biological interest does not encompass cases where markers all lie in lower-dimensional subspace or two or more of them coincide, which contrasts with Kendall's spaces, which contain the shapes of all possible configurations except those for which all the points coincide (Kendall *et al*., [1], p. 2).The fact that Kendall's shape-space is non-Euclidean, that is, curved, leads to difficulties in statistical appraisals. For triangles, the space is equivalent to the surface of a sphere of unit diameter (Figure 1), but for more than three landmarks, the space is of higher dimension and more complex to deal with. However, performing a multivariate statistical analysis in the tangent plane to Kendall's shape-space provides a valuable technique for circumventing the most serious of these difficulties. The manipulation involved is equivalent to the way in which a cartographer projects a map from a globe onto a flat sheet. The coordinates of the points representing the specimens are no longer in terms of the sphere, but rather are coordinates in the plane.

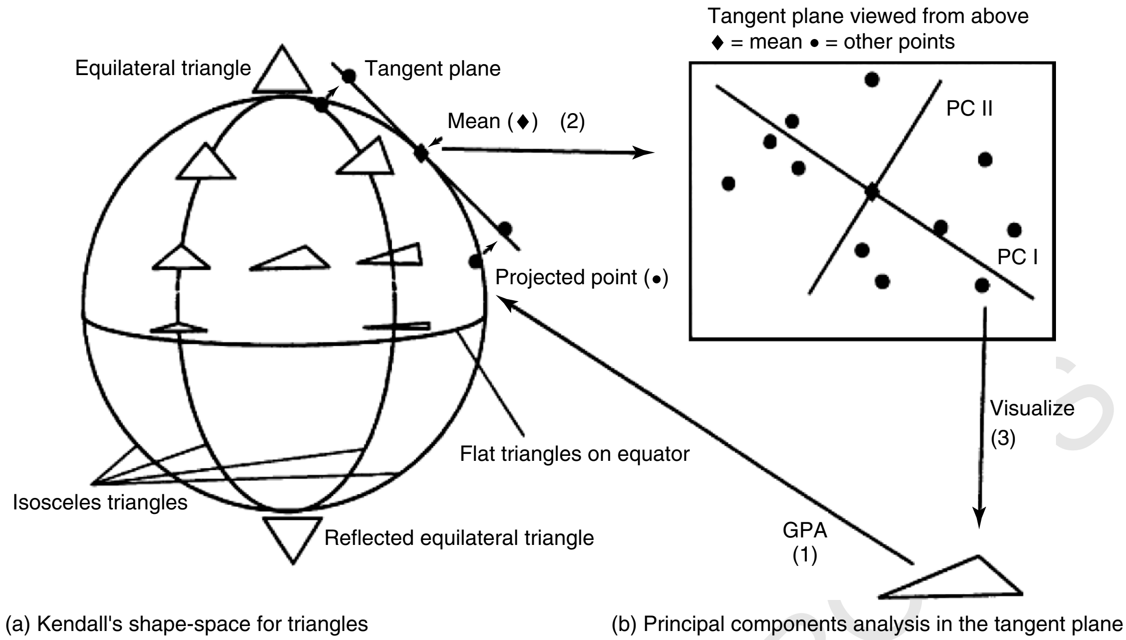

Figure 1 Kendall's shape-space for triangles. (Reproduced from Ref. 12. *Copyright* Wiley-Blackwell Sons, Anatomical Society of Great Britain and Ireland, 1998.)

This is an approximate solution but it works quite well provided the projection has not resulted in excessive distortion, such as would occur if the projection were to cover a large proportion of the sphere.

Deformation

Of the possible methods available for the study of deformations, the thin-plate spline method (*see* **Splines in nonparametric regression**) has taken pride of place [16, 17], a technique well known in physics. This function, which is applied to the whole landmark configuration, minimizes the bending energy of the deformation and hence results in minimal local variation of grid elements. The grids are interpreted as indicating how the space in the region of a reference shape might be deformed into that in the region of the target; that is, the landmarks of the reference are made to map into their counterparts in the target. The resulting visualization is very appealing but it should be noted that the manipulation is mathematical and is based on the locations of a few points, the homology or equivalence of which is known beforehand. The thin-plate spline interpolation is useful for exposing nonaffine and affine components of shape differentiation and for exploring the variation in different scales with respect to local variations versus global variations. A development by Bookstein [9] applies the catastrophe theory of Thom for extracting spatially discrete, localized details of transformation grids, thereby enhancing the interpretability of warp visualizations in a biological frame of reference.

An appropriate statistical procedure for linking shape variation (deformations) to the influence of factors in the environment (*see* **Cross-scale morphology**) is due to F.J. Rohlf and M. Corti. Adams and Rohlf [18] analyze ecological character displacement for salamander species, using geometric morphometric methods, geographical location, and stomach contents. Significant morphological differences were found to be associated with the factor of prey consumption. Tabachnik and Bookstein [19] placed morphometric variation in planktic foraminifers in a general ecological context.

Software for carrying out geometric morphometric analysis is available from http://life.bio.sunysb.edu/morph. This is a service to the scientific community maintained by Professor F. James Rohlf.

Examples

The examples briefly outlined here give a general orientation to the three types of procedures above.

A representation of Kendall's shape-space for triangles is shown in Figure 1(a). Equilateral triangles lie at the poles; the southern hemisphere is a reflection of the northern. The sphere is subdivided into 12 equal half-lunes (six in each hemisphere). If the apices of the triangles are unlabeled and reflections are ignored, then all triangles lie in one half-lune. Isosceles triangles lie along the lines dividing lunes and flat triangles at the equator. Figure 1(b) is a schematic diagram indicating the projection of points representing triangles in Kendall's shape-space in a space tangent to the mean (arrows) and the principal components (PCs) of shape-variability (PCI, PCII) in this tangent space. The computational steps are (i) generalized Procrustean analysis to register figures, which are then represented as points in the shape-space; (ii) projection of points into a space tangent to the mean and the PCs of shape variation in this space are calculated; (iii) visualization of the shape variability represented by the PCs is realized by reconstructing figures.

An example of the Q-mode ordination by principal coordinate analysis for ecophenotypic variations in a Moroccan Cretaceous ostracod species is shown in Figure 2; Q-mode analyses focus on association between individual units. Photographs of the actual specimens of *Oertliella tarfayaensis* are shown superimposed on the plot. In this manner, the shape and ornament of each of the individuals may be compared and contrasted visually. The variables are six in number, all dichotomous or discontinuous, namely, the numbers of anterior and posterior spines, the nature of the lateral ornament (reticulated or not), form of the posterior (blunt or subcaudate), and structure of the lateroventral ornamental ridge.

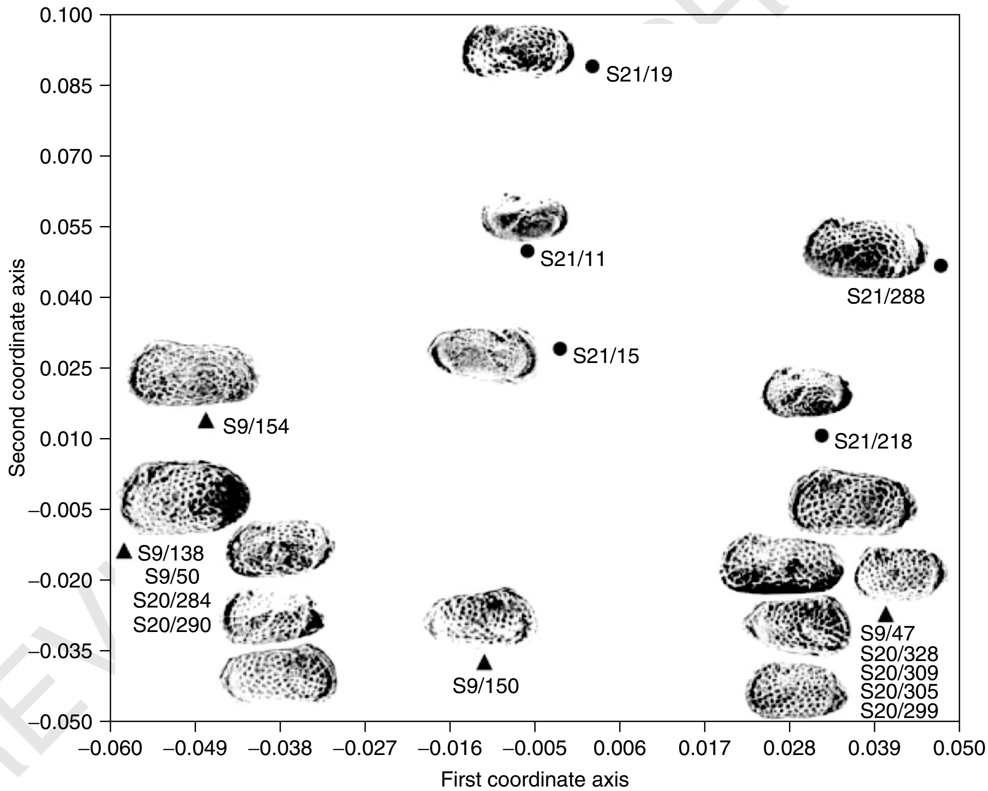

Figure 2 Standard multivariate analysis and ordination by ornamental features. (Reproduced from Ref. 20 by permission of the Micropaleontology Press, American Museum of Natural History, New York.)

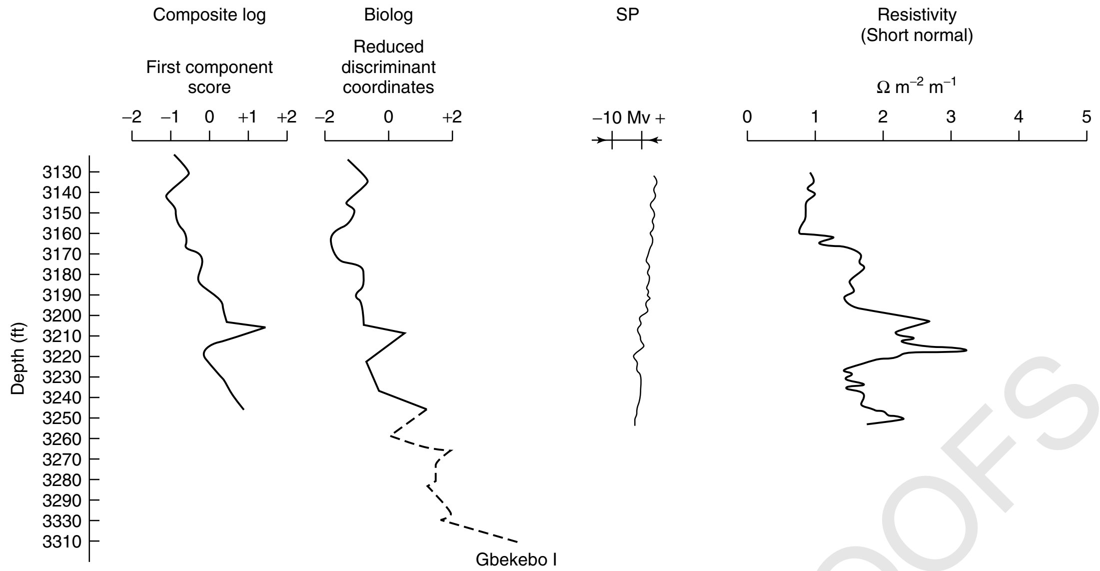

Figure 3 Burnaby's growth-free transformation for distance measures and paleoecological parameters. (Reproduced from Ref. 21 by permission of Elsevier. Copyright (1978) Elsevier.)

An application of Burnaby's [4] growth-invariant solution to physical environmental measures, a distance-measure example, is shown in Figure 3. The figure shows the juxtaposition of a growth-invariant Q-mode latent-vector (first vector) scores (labeled "biolog"), and electrical drilling logs, showing a

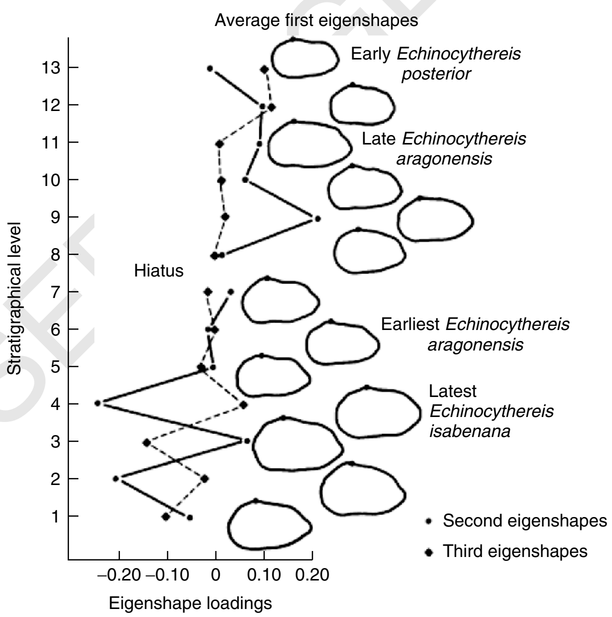

Figure 4 Outline morphometrics applied to an evolutionary sequence under environmental stress. (Figure based on data kindly provided by ELF, Pau, France.)

monotonic trend for physical characters and the shape descriptor. The data are for the Nigerian Cretaceous foraminiferal species *Afrobolivina afra* Reyment.

An example of outline morphometrics is shown in Figure 4. The figure shows the plot of "eigenshapes" for transitions between three species of the Eocene ostracod genus *Echinocythereis* from Aragon, Spain. The outline figures of left valves are average first eigenshapes. The pair of plots located to the left show the paths traced out by the second and third eigenshapes. The sequence sampled reflects environmental stress occurring in an area undergoing a long-term shift in sea level.

Figure 5 shows the ordination of specimens by the third relative warp plotted against centroid size for ecophenotypic variation in the carapace of the ostracod species *Hemicythere fulva* McKenzie, Reyment and Reyment from the Eocene of south-eastern Australia. The outlines of shells were made from stereoscan photographs. The ecophenotypy may be the outcome of a deficiency in calcium carbonate during the secretion of the shell. The inset deformation grid displays the principal warp comparison of the average morph "lace," an ecophenotypic variant, with the average morph "normal."

Comments

Dryden and Mardia [17] stress the practical aspects of shape analysis with arguments that should appeal to environmetricians wishing to apply shape-analytical procedures. Kendall *et al.* [15] is a substantial mathematical treatise on the theory of shape *s.l.*, widely conceived and in which biological questions are not a major issue. This work constitutes an inroad for the mathematician wishing to probe the structure of shape spaces. Mardia [11] gives a useful introductory account of statistical shape analysis and some of its applications. A recent and didactically useful account of shape statistics, superimpositions, and tangent spaces is that of Rohlf [23]. A caveat may be necessary: if environmetric data in the form of tables of percentages or tables of frequencies are to be joined in a multivariate statistical analysis with morphometric parameters, then the appropriate log-ratio statistical procedures for compositional data analysis should be adopted [21, 24]. A paleoecological problem occurring in the postmortem history of, for example, cephalopod shells concerns whether the shells occur *in situ* or were transported by water currents after death. Reyment [25] exemplified the

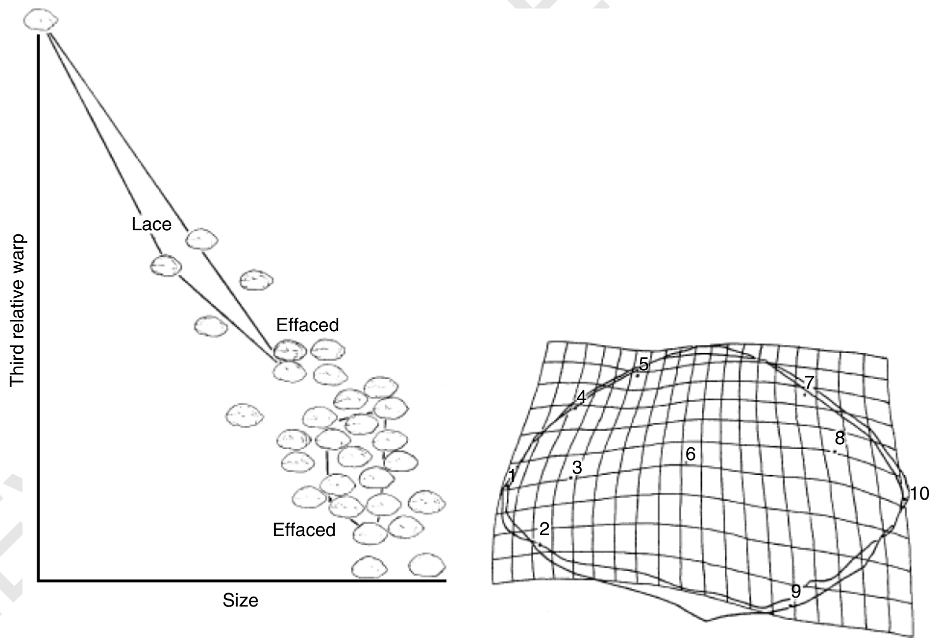

Figure 5 Geometric morphometrics and ecophenotypy in fossilostracods. (Reproduced from Ref. 22 by permission of *Revista Espanola de Paleontologia*.)

problems involved using shells of species of the ammonite genus *Knemiceras* from the Lower Cretaceous of Lebanon and Iran employing multinomial classification rules for discrete classification, in particular the Dillon–Goldstein Classification rule (Pires and Branco [26]). The Russian paleontologist J. A. Efremov coined the term *taphonomy* for the study of burial, sedimentation, and preservation of animals and plants. Reyment [25] introduced the term *taphonometrics* for the statistical study of such data.

References

[1] Rao, C.R. (1952). *Advanced Statistical Methods in Biometric Research*, John Wiley & Sons, Ltd, New York.

[2] Blackith, R.E. (1957). Polymorphism in some Australian locusts and grasshoppers, *Biometrics* **13**, 183–196.

[3] Blackith, R.E. & Albrecht, F.O. (1959). Morphometric differences between the eye-stripe polymorphs of the red locust, *Scientific Journal of the Royal College of Science* **27**, 13–27.

[4] Burnaby, T.P. (1966). Growth-invariant discriminant functions and generalized distances, *Biometrics* **22**, 96–110.

[5] Olson, E.C. & Miller, R.L. (1958). *Morphological Integration*, University of Chicago Press, Chicago.

[6] Lohmann, G.P. (1983). Eigenshape analysis of microfossils. A general morphometric procedure for describing changes in shape, *Mathematical Geology* **15**, 659–672.

[7] Thompson, D.W. (1961). *On Growth and Form*, Cambridge University Press, New York, Version of 1917 abridged by J.T. Bonner.

[8] Bookstein, F.L. (1991). *Morphometrics Tools for Landmark Data*, Cambridge University Press, New York.

[9] Bookstein, F.L. (1998). A hundred years of morphometrics, *Acta Zoologica Academiae Hungaricae* **44**, 7–59.

[10] Kendall, D.G. (1984). Shape-manifolds, Procrustean metrics, complex projective spaces, *Bulletin of the London Mathematical Society* **16**, 81–121.

[11] Mardia, K.V. (1999). Statistical shape analysis and its applications, in *On Growth and Form: Spatio-temporal Pattern Formation in Biology*, M.A.J. Chaplain, G.D. Singh, & J.C. McLachlan, eds, John Wiley & Sons, Ltd, New York, pp. 337–355.

[12] O'Higgins, P. & Jones, N. (1998). Facial growth in *Cercocebus torquatus:* an application of three-dimensional geometric morphometric techniques to the study of morphological variation, *Journal of Anatomy* **193**, 251–272.

[13] Reyment, R.A. (2010). Morphometrics: an historical essay, in *Morphometrics for Non-morphometricians*, Lecture notes on earth sciences, A.M.T. Elewa, ed., Springer Verlag, pp. 9–24.

[14] Reyment, R.A. (1991). *Multidimensional Palaeobiology*, Pergamon Press, Oxford.

[15] Kendall, D.G., Barden, D., & Le, H. (1999). *Shape and Shape Theory*, John Wiley & Sons, Ltd, New York.

[16] Bookstein, F.L. (1989). Principal warps; thin-plate splines and the decomposition of deformations, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**, 567–585.

[17] Dryden, I.L. & Mardia, K.V. (1998). *Statistical Shape Analysis*, John Wiley & Sons, Ltd, New York.

[18] Adams, D.C. & Rohlf, F.J. (2000). Ecological character displacement in *Plethodon:* biomechanical differences found from a geometric morphometric study, *Proceedings of the National Academy of Sciences* **97**, 4106–4111.

[19] Tabachnik, E. & Bookstein, F.L. (1990). The structure of individual variation in Miocene Globorotalia, DSDP site 593, *Evolution* **44**, 416–434.

[20] Reyment, R.A. (1978). Quantitative biostratigraphical analysis exemplified by Moroccan Cretaceous ostracods, *Micropaleontology* **24**, 24–43.

[21] Reyment, R.A. (1978). Graphical display of growth-free variation in the Cretaceous benthonic foraminifer *Afrobolivina afra*, *Palaeogeography, Palaeoclimatology, Palaeoecology* **25**, 267–276.

[22] Reyment, R.A. (1993). Ornamental and shape variation in Hemicytherura fulva McKenzie, Reyment and Reyment (Ostracoda, Eocene, Australia), *Revista Espanola de Paleontologia* **8**, 125–131.

[23] Rohlf, F.J. (1999). Shape statistics: procrustes superimpositions and tangent spaces, *Journal of Classification* **16**, 197–223.

[24] Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*, Monographs on statistics and applied probability, Chapman & Hall, London.

[25] Reyment, R.A. (2010). A discrete classification procedure applied to encrusted ammonite conches – a contribution to Taphonometrics, *Mathematical Geosciences* **42**, 447–452.

[26] Pires, A. & Branco, J. (1997). Comparison of multinomial classification rules, *Journal of Classification* **14**, 137–145.

(*See also* **Classification**; **Tree morphology**; **Principal components**)

RICHARD A. REYMENT

Please note that the abstract and keywords will not be included in the printed book, but are required for the online presentation of this book which will be published on Wiley's online publishing platform. If the abstract and keywords are not present below, please take this opportunity to add them now.

The abstract should be a short paragraph up to 200 words in length and keywords between 5 to 10 words.

Author affiliations will appear in a List of Contributors in the front of the print edition and will also appear in the online version of the encyclopedia. Affiliations may have been edited in accordance with the style of the encyclopedia. Please check that the affiliation(s) given below are correct.

Abstract: Morphometrics is defined as the application of multivariate statistical procedures to the study of variation in morphological characteristics of organisms. As originally implemented, the analyses are based on standard measures ("distances") spanning critical sites. A more recent introduction is concerned with quantifying outlines of organisms, the data then being a suite of arbitrary coordinate-pairs around the circumference of the object, obtained by some automatically operating digitization procedure (e.g., the Zahn-Rosskies method). Geometric morphometrics is the latest class to appear. It relies on diagnostically located "landmarks" that form the basis for superimposition of coordinate systems, visualizations by deformation grids. The first and third of these categories can, without particular difficulty, be allied with environmental indicators by appropriate techniques.

Keywords: morphometrics; multivariate statistics; environmetics; kendall's shape space; geometric morphometrics

Author Contact Address:

RICHARD A. REYMENT

Section for Palaeozoology, Swedish Museum of Natural History, Stockholm, Sweden