Lecture Notes in Earth Sciences 124

Morphometrics for Nonmorphometricians

Bearbeitet von
Ashraf M.T. Elewa

1. Auflage 2012. Taschenbuch. xi, 367 S. Paperback
ISBN 978 3 642 26370 5
Format (B x L): 15,5 x 23,5 cm
Gewicht: 581 g

Chapter 2
Morphometrics: An Historical Essay

Richard A. Reyment

Idea and Aims

The aim of this chapter is to inform morphometricians at large on the historical background of MORPHOMETRICS and to present the work and thoughts of the originators of the subject, with emphasis on the pioneering achievements of Professor Robert E. Blackith. Links between the latest developments of an aspect of morphometrics, in which geometric aspects are stressed, are outlined.

Introduction

The concept of morphometrics has a long history, notwithstanding that most of what is housed under its aegis is of quite recent origin. In this essay an outline of the two faces of morphometrics is attempted. The emphasis is on the historical development of the branch of knowledge over time, the orientation being one of respecting a state of intellectual achievement *à l'epoque* and without pointing the finger or scorn at past work considered proven by current thought as being misguided or inadequate. Reyment (2005) is a more detailed review of applied morphometrics in thought and praxis to which the reader may be referred for notes on palaeontological applications.

Morphometrics may be defined as a more or less interwoven set of largely statistical procedures for analysing variability in size and shape of organs and organisms. Some of the concepts have been generalized to encompass non-biological problems. Such areas are not taken up in the following. For a complete account of the subject of generalized shape as a mathematical concept, the most authoritative reference is Kendall et al. (1999). Attempts at expressing variability in size and shape in quantitative terms have a relatively long history in biology. The basic principle of what came to be known as principal component analysis was also presented

---

R.A. Reyment (✉)
Naturhistoriska Riksmuset, S10405, Stockholm, Sweden
e-mail: richard.reyment@nrm.se

A.M.T. Elewa (ed.), *Morphometrics for Nonmorphometricians*, Lecture Notes in Earth Sciences 124, DOI 10.1007/978-3-540-95853-6_2, © Springer-Verlag Berlin Heidelberg 2010

in a shape-size context. However, its development and application were held back owing to computational difficulties. Some noteworthy biological studies did, however, emerge for material based on just a few variables, for example, the work of Pearce (1959, 1965) on the expression of shape in fruit-trees.

The term "Morphometrics" seems to have been coined by Robert E. Blackith (Professor of Zoology at University College, Dublin University) some 50 years ago (Blackith 1957), who entered into the subject through his engagement with the agricultural problem caused by swarming locusts. He 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. This approach is clearly biological and Blackith's results represent one way of introducing a precise biological model into a statistical analysis of traditional stamp.

Essential to morphometric analysis is the requirement that variation in shape be represented in describable and repeatable terms. The roots of the geometric aspects of morphological graphology is usually conceded to lie with the work of the 15th century artist Albrecht Dürer, usually said to be German but who in actual fact was Hungarian. Dürer's father relocated to Nüremberg where he changed his name first to Thürer (from the Hungarian Ajtósi, doormaker), then to Dürer, more in keeping with the local dialectal pronunciation of Thürer. In geometrical terms, Dürer used the properties of affine transformations for distorting details (such as faces) by lateral or vertical elongation, thus mapping a square into a parallelogram and, or, by shearing a feature.

Blackith (1965) pointed out that the earliest recorded attempts to compare the shapes of animals was made by the school around Pythagoras as early as the 5th century BC. A rough drawing of a plant or animal that recorded the essence of the organism by noting the number of junctions between the lines of the sketch, but with no more than the minimum number of lines for representing the image of the organism provided the basis of the procedure.

Somewhat later it seems, the ancient Egyptians were concerned with embellishing burial monuments with figures and scenes, carved in limestone. Examination of photographs of some of these works of art, for example those figured in the treatise by Gay Robins (1994) discloses that some of them retain a discernible pattern of standardized squares, marked in red "chalk" which form a framework for making the carvings. It is apparent that the finished work of art was meant to be cleansed of the "props", but this did not always take place. How were the squares useful? She, and others, have noted that there was a reigning system of conventions, presumably determined by a collegium of experts, which defined the proportions of the human body to be used in adorning graves and memorials. The proportions of the limbs were standardized to a given number of squares, or a part of a square. It is only in the details of the head that a genuine likeness could be introduced. The square-standards were maintained for hundreds of years, then changed for a lengthy period, only to be brought back to the original conventions, probably by decree.

The skilfully stylized system of the ancient Egyptians for representing human proportions was to appear many hundreds of years later in the hands of Dürer and da Vinci, and at a time when the system of rules of several thousand years earlier had neither been understood nor, even, suspected. Dürer's "menschliche Proportionen" were little more than the standardized Egyptian squares, though now expressed as lengths in terms of fractions of total height and, at times, affinely transformed. The well known figure of a man inscribed in a circle attributed to Leonardo da Vinci is a demonstration of the same idea, more succinctly assessed by Dürer (1512, 1513, 1520).

The origins of the quantitative analysis of measurements on organisms lie with the early biometricians, Karl Pearson (1857–1936), Francis Galton (1822–1911) and W. F. R. Weldon (1860–1906). A review of the pioneer studies of this triumvirate, and others), is given in (Reyment 1996). One of the methods, the correlation coefficient, invented by Galton and mathematically anchored by Pearson, is a standard tool of biometric analysis. It is worth considering why Pearson did not make more of his finding that correlations computed between proportions are not valid – he called them spurious correlations. The reason may well be, as has been surmised, that Pearson was not really interested in becoming entangled in biological conjecture, his main interests lying elsewhere. Nonetheless, two years before his death, Pearson and Morant (1934) published a detailed biometrical analysis of the "Wilkinson skull", proving beyond doubt that it was indeed the mummified skull of Oliver Cromwell, left to its fate on a pole after a gruesome post-mortem beheading.

Morphometrics as an Algebraic Exercise

As effective computational facilities became more generally available, so did interest in studying shape variation increase. A rather simple procedure rapidly gained ascendancy that relies on interpreting the latent roots and vectors of a matrix of covariances or correlations in terms size and shape components. The idea seems to have been formulated by the French marine biologist Teissier (1938) in his work on crabs (but only for the first principal component) and then improved and introduced into the anglophone literature by Jolicoeur and Mosimann (1960) and later provided with a more logical mathematical model by Hopkins (1966) using principal component factor analysis. Jolicoeur (1963) developed his original idea further by successfully relating the latent vector reification of principal component analysis to an allometric model. Although the general "algebraic solution" is usually attributed to Jolicoeur and Mosimann (1960), Quenouille (1952) was, it seems, probably the first worker in the field to give latent vectors a biological shape-oriented interpretation, closely followed by Pearce, who summarized his work on the growth of trees in a statistical textbook (Pearce, 1965). Sprent (1972) proposed a formalization of the prevailing concepts at the time for the analysis of size and shape.

The principal component method for describing size and shape variation relies on an intrinsic property of non-negative matrices known as the Perron-Frobenius theorem which states that the maximum root of the matrix is associated with a latent vector with positive components. The first latent vector of a covariance or correlation matrix of distance-measures observed on some organism is interpreted as indicating variation in size. Subsequent latent vectors are said to be indicative of various factors of shape-variation. This interpretation is circumstantial in that it is in part based on an artefact. Nevertheless, the method is still in wide use and there is some evidence that supports the claims for a useful spectral decomposition of variability in size and shape. The same methodology is applied in Geology to the study of sedimentary data, species compositions in palaeoecology, and in the analysis of geochemical data. In these connexions there may be more convincing justification for the interpretation (reification) of the latent vectors. The results of multivariate analyses of distances observed on fossils are often presented as ordinations. That is, as bivariate scatter plots of the scores obtained from substituting the columns of the data-matrix into the latent vectors.

Growth and shape-change include a tensorial component in that they are at different rates in different directions at typical locations in a tissue. Huxley (1932) was well aware of this distinction. This directionality is not easy to test in a principal component decomposition, but easy of access if a geometrical orientation is adopted.

Mosimann's (1970) paper on allometry, and the " identification" of the size vector can be said to mark the starting point for a more geometrical approach to morphometrics and one which lies near to the heart of the new geometric morphometrics but without going all the way.

Blackith (1957) embarked upon an ambitious programme aimed at establishing a quantitative basis for studying growth and form in insects. The initial publication can be seen as an attempt at mapping out a program for detailed research with standard methods of multivariate analysis such as were available at the time, in a biological vista and with emphasis on the results being achieved by the Indian school (Calcutta) in the hands of P. C. Mahalanobis and C. R. Rao (1952) with reference to the anthropometrical survey of the cast- system in India (United Provinces). The procedures reviewed in that work were the linear discriminant function, the method of multiple discrimination (canonical variate analysis) and principal components. A feature of this short article is the use made of a chart of generalized distances along two axes, one for phase and one for size. Within the field of experimental biology, Blackith (1957, 1960), in a suite of papers devoted to polymorphism in insects, in which Mahanolobis' distances and canonical variates were handled with great insight, established a foundation for the quantitative study of variation in size and shape. A major problem to be solved in using "distance measures" in the study of variation lies with the fact that size and shape are confounded. This means that a suite of measurements on a locust carapace, for instance, represent not only a component expressing size-differences but also differentiation between variables that arises from differences in shape of the specimens of a sample. Blackith could reduce the negative aspect of this by a clever graphical method of linking his distances in relation to known polymorphisms. This could not however hope to provide a general, unequivocal solution to the problem of the analysis of shape.

Blackith et al. (1963, p. 317) observed that which has also been done by others that Thompson's (1917) application of coordinate transformations was done in a semi-quantitative manner and never expanded and could not take account of (nor indeed was envisaged as doing) such a fundamental attribute as that of differential growth and the change or relative proportions with absolute size. A solution to that problem came first with Huxley (1932) in his celebrated volume on relative growth (N.B. Huxley did not use the term "allometry" in his treatise of 1932 for relative growth – his term was heterogeny).

The Q-mode method of principal coordinate analysis has shown itself to be useful for presenting graphical analyses of relationships such as specific differences and intraspecific differences such as polymorphism in characters, not only morphological features but also features that are dichotomous and qualitative. Gower (1966) developed this technique as an answer to the inappropriate analysis known as Q-mode factor analysis, a procedure which, however, is neither an authentic factor model nor constructed for preserving distances between specimens (Reyment and Jöreskog 1999). Principal coordinate analysis is, moreover, not inverted principal components. In order for the technique to function in a correct manner, Gower's (1971) similarity matrix is a necessary prerequisite. This use of this technique permits the user to combine quantitative, binomial and qualitative attributes in the same distance-preserving representation. In the special case of a matrix of correlations of correlation coefficients, principal components is the R-mode dual of Q-mode coordinates, and the necessary manipulations can be simply performed by Sylvester's extraction of latent roots and vectors – now generally known as the singular value derived by decomposition of a square symmetric matrix in the more general Eckart-Young solution. Be it noted however that this is a restrictive procedure in that it cannot encompass dichotomous and qualitative variables and hence is of limited value in very many studies, and particularly in palaeontology where the number of characters may be diverse as well as numerous.

How biologically sound are distance-based characters in the hands of mathematically oriented biologists? This is a question raised occasionally by the more geometrically schooled practitioners. My experience is that the selection of "taxonomically relevant distances", which are, in effect, distances between "landmarks", is usually done in a very conscientious manner and well anchored in a detailed knowledge of the biology of the species under study. The proof of the pudding is in the eating, as it were, and in the case of entomological work on harmful insects, it is been found by experience that distance measures, judiciously selected, yield valid, intelligible results. The term "landmark" is not a great terminological borrowing. Its invasion of the biometrical literature comes via its applications in craniology, the study of variation in the dimensions of skulls of primates, and basically a subject not of prime interest to the mathematician. The real meaning of the word is, as any anglophone knows, (1) object marking boundary of country, estate, etc., (2) conspicuous object in a district, (3), object or event or change marking stage or process or turning-point in history.

Before leaving the insects, it is interesting to note that Blackith and Kevan (1967, p. 81) took up the geometrical problem of the space in which canonical variates are located. Campbell (1979, Canonical variate analysis: some practical aspects. PhD Thesis, unpublished) has made an extensive study of the structure of the canonical variate model. He showed that it is not to be expected that the vectors of canonical variate analysis are directly equatable with the latent roots of principal component analysis. This is often a matter of confusion to the mathematically untrained users of computer programs.

Morphometrics as Interpreted by Statistical Mathematicians

The English mathematician Maurice S. Bartlett, who was intimately connected with the development of several major areas of mathematical statistics, also made contributions to the application of multivariate statistics to quantitative biology (Bartlett, 1965). Admirable though as this is as to the clear presentation of multivariate methodology, Bartlett's insight into biological processes leaves much to be desired.

T. P. Burnaby began his professional life as a palaeontologist at the University of Keele, U. K. His studies were from the outset centred around quantitative aspects of various invertebrates, mainly foraminifers and bivalves. Burnaby died prematurely (1924–1968) and his widow deposited his research notes with me at the University of Uppsala in 1970. These notes show Burnaby to have been gifted with mathematical acuity and it is clear that at the time of his passing he was concerned with several biological problems which, if they had been pursued to fruition, could well have led to major contributions in the analysis of growth and form. Burnaby was much concerned with arriving at a reliable procedure for studying size and shape in bivalves. Bivalves do not have a terminal growth-size and grow additively, which implies that just keep getting larger and larger right unto the end. Burnaby observed that the growth pattern was not a mere regular amplification of the shell but that there were steps that were lacking in perceptible regularity. He made a similar observation for the growth of coiled ammonite shells (genus *Cadoceras)* which on the basis of the Reverend Mosely's identification of the logarithmic growth spiral for ammonites, were considered to develop in a regular manner, but which did have a terminal growth-size. Burnaby's notes, and Burnaby (1966a), show that he had found that the angle of growth (the spiral angle) of presumably regularly coiled ammonite shells is not constant but can, and does, fluctuate during growth.

Thus, the real nature of the problem of confounding of size and shape during growth was first realized clearly by Burnaby (1966b) who devised a mathematical procedure for a transformation that placed "size" in one subspace and shape into another. He named his method "growth-free discrimination". This solution is not often used. It is nonetheless a remarkable and foresighted achievement that doubtlessly influenced directly or indirectly the current phase of development of the subject. Burnaby's (1966b) quest for a discriminant function that would override the effects of confounding due to size differences, polymorphism, ecologically stimulated effects (ecophenotypy), etc., representable as gradients, had a definite circumscribed objective. He was definitely not concerned with describing shape variability itself, as is sometimes assumed, but solely with addressing a biostratigraphical problem occurring with organisms that do not have a terminal growth size. Burnaby was fortunate in having C. Radakrishna Rao as one of the referees for his submission to *Biometrics* and for Rao being willing to help provide the text with a rigid mathematical framework (Burnaby's correspondence). A major problem at the time was that Burnaby was unable to find real data for exemplifying his procedure and was obliged to use an artificial example which almost led to the rejection of the submission (correspondence in Burnaby's manuscripts). Rao (1966) moved quickly to produce a theoretical paper based on his knowledge of Burnaby's submission and in part inspired by it. However, Rao concentrated on discrimination aspects and hence on two populations, but did not discuss problems of estimation.

Gower (1976) took up the algebra of Burnaby's growth invariant discriminant functions in a more detailed manner. Gower noted that the conceptual basis of the original work of Burnaby was not unchallengeable, not least because of the problem associated with the estimation of linear growth effects and a difficulty which Burnaby had not been able to surmount. Reyment and Banfield (1976) applied Gower's panoply of methods for growth-adjusted canonical variates to species of Paleocene foraminifers. Among the methods proposed by Gower for estimating growth etc. vectors, we used two of them, to wit, the principal component solution and approximation by maximum likelihood factor analysis (i.e. "True factor analysis" (Reyment and Jöreskog 1999). The principal aim of our analyses was to use the canonical variate means for comparisons between species observed at different time-intervals and not to attempt an analysis of shape variability. This is an important limitation on the scope of the study and one worth keeping in mind. The results obtained for the methods proposed by Gower (1976) showed that the principal components of the pooled within-samples covariance matrix were successful in removing the major source of conflicting variation. This variation was found to derive from individuals of the foraminiferal species having been at different stages of growth when fossilized. We also found that the patterns produced when the growth-free canonical variate sample means were plotted against chronological order led to useful evolutionary comparisons between species.

Supplemental Reading

The book by the plant-ecologist P. Greig-Smith (1957) is essential reading for anybody seriously interested in applying quantitative methods to the study of plants not least because of the detailed practical treatment of principal components and principal component factor analysis.

A becharming advanced introduction to biometric analysis is the book by C. R. Rao (1952), notwithstanding its almost 60 years since the first printing. Questions being asked by tyros today with respect to aspects of canonical variate analysis (Rao does not use this term in his book, preferring the designation multiple discrimination) are taken up in, for example, the case of Miss M. M. Barnard's (1935) Egyptian skulls and Fishers statistical treatment of her problem. And, as an intellectual exercise, the astute student is invited to find out what is wrong with Barnard's model from the biological aspect.

A mathematically rigorous treatment of what is known as "factor analysis" in many connexions is the text by Reyment and Jöreskog 1999. The distinction brought to the fore here is that the true factor model is seldom, if ever, encountered in biology. A surrogate of somewhat rickety validity is employed instead, the appropriate term for which is "principal component factor analysis". The confusion seems to have arisen with Teissier's (1938) expression "analyse factorielle" for a principal component analysis of crabs and a too literal translation, of the French text.

Reyment (1996) gave an historical review of the early background of morphometrics in which the evolution in thought on that subject is outlined.

The Coefficient of Racial Likeness was introduced into anthropometry by Karl Pearson. It was, however, nothing more than the generalized statistical distance of P. C. Mahanalobis (shown to Pearson in 1927 by Mahanalobis), with all off diagonal elements zero and variances along the diagonal. Pearson did not want to admit the superiority and relevance of the generalized distance, thereby embarking upon a fruitless controversy (cf. Mahanolobis, 1936).

Morphometrics in a Geometric Setting

The problem of how to construct a practical and mathematically justifiable solution to the coordinate-based concept of Thompson (1917, 1942) for geometrically illustrating shape relationships in organisms with the emphasis placed on phylogenetic reconstructions dogged biometricians for many decades. Thompson thought it should be possible to represent relationships between related organisms by deformational grids. By means of an affine transformation, one genus of fishes, for example, could be transformed graphically to another genus, hopefully phylogenetically connected. For years workers puzzled over how this was actually done – what did the mathematics look like? Speculation went on for years and years without much light being cast on the subject. Bookstein (1991) and Dryden and Mardia (1998, p. 200) noted the subjectivity attaching to Thompson's freehand transformations. Huxley (1932, pp. 104–110) discussed weaknesses he considered to be inherent in the method of Cartesian coordinates from the aspect of the constancy of growth-gradients. Thompson's book was reprinted many times after 1917 but he did not in any of the editions refer to how he had made his figures. In any event, Thompson's insight was a major break-through in thought about biological variability but it was not until 1978 that Bookstein provided a solution for the affine case and then for the non-affine case (Bookstein 1986) in many later publications, summarized in Bookstein (1989, 1991). These works lie centrally located at the origin of the modern development of geometric morphometrics. It is no exaggeration to claim that without Bookstein's exceptional insight, and didactic skills, the geometric analysis of shape would not have developed into a biologically relevant discipline, capable of being expanded in many directions.

Measuring Outline Shape

Distance measures can be arbitrarily constructed around an outline by a simple geometrical method by marking off equispaced points. A well known means of describing a curve, such as an outline, is by decomposing the line into a Fourier Series which can be made to approximate the contour of the object by passing through a series of progressively more complex trigonometric functions for the digitized points. Although very good approximations of the shape of an object can be made, the results cannot be linked to homologous relationships between objects. The use of Fourier Series is well known to geologists from the sphere of analytical sedimentology.

Lohmann (1983) applied a result of Zahn and Rosskies (1972) to the study of shape in planktic foraminifers. This result shows that a complex curve can be expressed as a series of steps around a circle. The underlying mathematical concept of Lohmann's "eigenshapes" is conceptually similar to that of Burnaby's decomposition into size and shape spaces. In eigenshapes, size is represented by the length of the steps around the perimeter of the object and shape by the set of angles estimating the deviation of each step from the expected direction. Lohmann's work, considered in the context of time and knowledge, is an outstanding achievement. Swiderski et al. (2002) have given a thoughtful appraisal of eigenshape analysis. Macleod's (2002) extended eigenshape analysis is a serious attempt at coming to grips with the problem of comparability between objects. Bookstein (1997) proposed powerful coordinate-based methods for studying forms without landmarks, that is, for accessing the information in curving outlines.

A reasonable case for outline methods can be made for planispirally coiled shells such as those of ammonites. Natural, circumferentially located reference points are not available on such shells, even where ornamental features occur, owing to the inherent instability in such properties with respect to the number per whorl and the degree of development. Ammonites are often richly ornamented with tubercles and ribs. These features are, however, seldom stable enough to permit using them as a base for homologous landmark points. Greater availability is offered by the apertural aspect in which points of intersection of features occur. Here again the canon of geometric morphometrics is not always useful and for many studies, the analyst must perforce fall back on the multivariate analysis of distance measures (Dryden and Mardia, 1998).

Reference Points (Landmarks)

Geometrically based morphometrics is in its current form dependent on the selection of reference points, designated by X-, Y-coordinates and conveniently referred to as "landmarks". MacLeod (2002) has drawn attention to this and, for example, pointed out that taxonomic distances in current use are no more than measures between reference points on an object which in turn are just "landmarks". The same distinction has been made by Dryden and Mardia (1998) and Bookstein (1991). Kendall et al. (1999, p. 1) espouse a rather rigid interpretation of labelled points (they adroitly bypass the use of the word "landmark") in their mathematical concept of markers in that they underline that for them, labelled 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-dimensional 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. 1999, p. 2).

Landmarks are specified by pairs of X-, Y-coordinates (the usage is originally a borrowing by osteologists from topographical surveying, where fixes are located with respect to coordinate pairs). Using simple geometry, distances can be constructed from taxonomic reference points: the reverse procedure is, however, not possible. The arbitrary points on a circumference in eigenshapes are likewise landmarks, denoted by X-, Y-coordinates, but they lack the property of homology or point-to-point correspondence from specimen to specimen.

Thin-Plate Spline-Based Morphometrics

The concept of thin-plate splines for expressing, graphically, deformations is not based on a biological concept, in common with the algebraic solution. It seems to have been deveoped in connexion with French engineering work, possibly connected with stability in fuselages of ultrasonic aircraft . One must be sure of what the spline-based morphometrics can do. Evangelical claims for biological relevance made especially in the earlier stages of applications have not always been useful and, as just stated, not true. The selection of landmarks determines the outcome of the shape-analytical conclusions owing to the fact that the thin-plate spline decomposition is not rooted in covariation in shape-changes in the input-data. Hence, each configuration of landmarks is a unique representation of just the set of points selected on the object (Rohlf, 2002, p. 179). Deletion or addition of points is often found to change the visualization diagram and hence the interpretation. In consequence hereof, it is desirable to make clear that the results of a particular analysis pertain specifically to a particular configuration of reference points on an organism. Biological interpretation must inevitably be made with expert insight (Rohlf, 1993).

Palaeontological material consists almost exclusively of fossilized hard parts, the shells encasing the tissues and organs of the animal. Material obtained from living organisms, for example, brains, can be related to more than raw morphology which implies that an analysis of an organ can be given a more biological interpretation than can be hoped for by a palaeontologist. For example Bookstein's (1991) work on schizophrenia. The study of shape-variability in fossilized hard parts must perforce be in terms of "deformations"; this limitation implies that exceptional care must be expended on choosing reference points that really mean something in a palaeobiological context and which can be extrapolated from case to case. In many cases, apart from fossil mammals, it is seldom possible to relate shape-variation in shell details to the anatomy of the soft parts of an animal. It is tempting to view a sequence of spline relationships resulting from the latent-root decomposition as being a kind of microscope-ratchet which successively yields a scale of magnifications of the surface of a fossil as a function of the latent-roots. For different configurations, different impressions of the topology of the surface will be communicated. Hence, there is considerable obligation placed on the investigator to select reference points that are of scientific significance. Large latent roots correspond to latent vectors that describe small-scale features - the deformation of landmarks that are close together. Small latent roots correspond to latent vectors that describe large-scale deformational features.

There is an important field in invertebrate palaeontology, however, that can supply much useful information from morphometric appraisal of shell properties and shape. This is the subject of polymorphism and polyphenism. Interesting results have been obtained for ostracods and cephalopods, for example, Reyment and Kennedy (1990) for ammonites and Reyment and McKenzie (1993) for Ostracoda from southern Australia.

Remarks on Spline-Based Methods

There are two main openings available for charting differences in form by means of coordinates. One of these takes each form, superimposes it in relation to others, and then computes differences in terms of reference-point displacements relative to this registration. The second tack is concerned with describing differences in point configurations as deformations of a grid produced by mapping one form into the other and visualizing the shrinkings and stretchings that are generated by the procedure. The analysis of the registrations of the coordinates may be done in several ways. 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 is usually referred to as generalized Procrustean fitting. The result is scaled by division by the centroid size.

Affine and Non-Affine Transformations

The concept of affine transformations seems to be due to the celebrated mathematicians Möbius and Euler. In Physics, the idea of affine transformation is known as an homogeneous deformation (Klein, 1925, p. 75). Decomposition of coordinate-based data by the thin-plate spline technique exhibits a close analogy with the well known decomposition by means of Fourier trigonometric series. The constant term in such a series is a global parameter and the trigonometric coefficients are local parameters at successively smaller scales (Dryden and Mardia, 1998, p. 199). Dryden and Mardia (1998, p. 286) also point out that the results yielded by distances and coordinates are often similar and the former cannot be dismissed out of hand. For small variations, registration methods and distance methods lead to identical conclusions about shape. Unfortunately, there is no general rule available to support this belief.

Possible Sources of Error

Errors can and do arise in the multivariate analytical processing of data generated by geometric morphometric procedures. The intuitively attractive method of triangulating data, Bookstein Coordinates, whereby coordinates are registered on a common edge induces spurious correlations and consequently invalid covariance matrices (Dryden and Mardia, (1998, p. 173). This notwithstanding, the basic theorem that everything that can be achieved using shape-coordinates can also be realized by means of ratios of distances, as demonstrated in the original paper on shape coordinates (Bookstein, 1986). The possibility of applying the theory of compositional data analysis (Aitchison, 1986) to the problem has not yet been more than briefly considered. Another source of error lies with imperfections in the data. This is more likely to be a problem in palaeontological work where deviations from multivariate normality are a cause of misleading results. Campbell (1979, Canonical variate analysis: some practical aspects. PhD Thesis, unpublished) seems to be the first worker to have given serious attention to practical aspects of consistency and repeatability in multivariate analysis, in particular canonical variates and discriminant functions, but also principal components. Some of Campbells's results are summarized in Reyment et al. (1984) and Reyment (1991).

Supplementary Reading

R. H. Benson (1972): Benson, a specialist on fossil and extant ostracods, was concerned with features of shell morphology, explicitly the reticulate pattern which he believed was under the control of evolutionary and environmental factors. This theory is not unchallengeable but it does contain an element of bio-geometric interest. He devised a method of graphical pattern analysis to be used for defining elements of the reticulum in a lavishly illustrated monograph. Benson's seldom cited monograph features series of reticular silhouettes with homologous regions identified with exemplary insight. The data were not analysed statistically, but remained at the visual level. Arising from Benson's graphical analysis, Siegel and Benson (1982) proposed a method for comparing shapes.

N. A. Campbell's (1982) publications listed in the references. These publications treat the problem of achieving stable results in methods of multivariate analysis, primarily canonical variate analysis but also principal components, with emphasis on the stability of the elements of the vector components. This is a question of more than trivial significance when the results obtained for, say, relative warps are to be expressed graphically or to be inserted into a standard extension of multivariate techniques. As far as I have been able to judge, this awareness has yet to make itself apparent in geometric morphometrics.

Bookstein's (1978) publication on interpreting affine transformations has made a lasting impression on me and one that impresses by the depth of understanding it contains. Every student of morphometrics should consult that thesis now and then. The generalization of the field to encompass affine and non-affine changes in shape by Bookstein (1986) is likewise a classic. Both papers mark the dawning of the new era for the study of shape and growth.

A very important advance in the analysis of shape-data is the paper by Mardia and Dryden (1989). This work may be seen as marking the beginning of the use of statistical models for shape studies. The term "Mardia-Dryden" is due to Kendall (1991).

Concluding Remarks

The brief account of morphometrics presented in the foregoing pages can do no more than highlight a few of the most important features and events of what is a complicated subject caught up in a phase of expansive development. The introduction of advanced geometrical thinking into statistics is not widespread and very few statistical textbooks take it into consideration. The level of mathematical theory involved in shape-theory is certainly beyond the reach and learning of many statisticians and certainly transcends the ability of geologists and biologists. The truth of this statement can be assessed by consulting the volume by Kendall et al. (1999). An interesting and promising development in biological shape analysis has been proposed by Bookstein (2000) by means of a method called "creases", being an allusion to the pinched features characterizing the associated spline diagrams. The method of crease analysis examines the effect of expansions forwards and backwards in time in phylogenetic reconstructions (Bookstein, 2000). As far as I am aware, the potential of the 'crease' has yet to be exploited, which may reflect a dearth of phylogenetic analyses at the present time.

The crease method also holds promise of being a means of developing the automated description of diagnostic contrasts between reference specimens of species, that is, image-based taxonomy.

Acknowledgments I wish to thank the Trustees of the Swedish Museum of Natural History, Stockholm, for providing working facilities. I express my gratitude to Professor Fred L. Bookstein for generously reading and constructively commenting on the manuscript and to Professor Kanti Mardia for valuable advice.

References

Aitchison J (1986) The statistical analysis of compositional data, Chapman and Hall, London, xv, 416 pp

Barnard MM (1935) The secular variation of skull characters in four series of Egyptian skulls. Annals of Eugenics (London) 7: 89 pp

Bartlett MS (1965) Multivariate statistics. In Waterman TH, Horowitz HJ (Eds.) Theoretical and mathematical biology, Chapter 8, Blaisdell Publishing Company, Newyork, 201–223 pp

Benson, RH (1972) The *Bradleya* problem, with descriptions of two new psychospheric ostracode genera, *Agrenocythere* and *Poseidonamicus* (Ostracoda, Crustacea). Smithsonian Contributions to Paleobiology, 12, 138 pp

Blackith RE (1957) Polymorphism in some Australian locusts and grasshoppers. Biometrics 13: 183–196

Blackith RE (1960) A synthesis of multivariate techniques to distinguish patterns of growth in grasshoppers. Biometrics 16: 28–40

Blackith RE (1965) Morphometrics. In Waterman TH, Horowitz HJ (Eds.) Theoretical and Mathematical Biology, Chapter 9, Blaisdell Publishing Company, Newyork, 225–249 pp

Blackith RE and Kevan DK (1967) A study of the genus *Chrotogronus* (Orthoptera). Evolution 21, 76–84

Blackith RE, Davies RG, Moy EA (1963) A biometric analysis of development in *Dysdercus fasciatus* (Hemiptera: Pyrrhocaridae). Growth 27: 317–334

Bookstein FL (1978) The measurement of biological shape and shape change. Lecture Notes in Biomathematics.– No. 24, Springer Verlag, New York, 191 pp

Bookstein FL (1986) Size and shape spaces for landmark data in two dimensions. Statistical Science 1: 181–242

Bookstein FL (1989) Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions Pattern Analysis and Machine Intelligence 11: 567–585

Bookstein FL (1991) Morphometric tools for landmark data. Cambridge University Press, Cambridge, xvii, 435 pp

Bookstein FL (1997) Landmark methods for forms without landmarks: localizing group differences in outline shape. Medical Image Analysis 1: 225–243

Bookstein FL (2000) Creases as local features of deformation grids. Medical Image Analysis 4: 93–110

Burnaby TP (1966a) Allometric growth of ammonite shells: a generalization of the logarithmic spiral. Nature 209: 904–906

Burnaby TP (1966b) Growth-invariant discriminant functions and generalized distances. Biometrics 22: 96–110

Campbell NA (1982). Robust procedures in multivariate analysis. II Robust canonical variate analysis. Applied Statistics 31: 1–8

Dürer, A (1512, 1513, 1525) Vier Bücher von menschlicher Proportion. Nuremberg, Hieronymus Formschnyder, 264 pp

Dryden IL, Mardia KV (1998) Statistical shape analysis. Wiley & Sons, New York, xviii + 347 pp

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrics 53: 12–38

Gower JC (1971) A general coefficient of similarity and some of its properties. Biometrics 27: 857–874

Gower JC (1976) Growth-free canonical variates and generalized distances. Bulletin of the Geological Institutions of the University of Uppsala. N.S. 7: 1–10

Greig-Smith P (1957) Quantitative plant ecology. Butterworth, London, 256 pp

Hopkins JW (1966) Some considerations in multivariate allometry. Biometrics 22:747–760

Huxley JS (1932) Problems of relative growth. Methuen, London, xix, 276 pp

Jolicoeur P (1963) The degree of generality of robustness in *Martes americana*. Growth 27: 1–27

Jolicoeur P, Mosimann JE (1960) Size and shape variation in the Painted Turtle, a principal component analysis. Growth 24: 339–354

Kendall DG (1991) The Mardia-Dryden distribution for triangles – a stochastic calculus approach. Journal of Applied Probability 28: 225–230

Kendall DG, Barden D, Carne TK, Le H (1999) Shape and shape theory. Wiley and Sons, Chichester, x, 306 pp

Klein F (1925) Elementarmathematik vom höheren Standpunkte aus. Band 2: Geometrie. J. Springer Verlag, Berlin, 302 pp

Lohmann GP (1983) Eigenshape analysis of microfossils: a general morphometric procedure for describing changes in shape. Mathematical Geology 15: 659–672

Macleod N (2002) Phylogenetic signals in morphometric data. In MacLeod N, Forey PL (Eds.) Morphology, shape and phylogeny; Systematics association special volume 64, Taylor and Francis, London, 100–138 pp

Mahanolobis PC (1936) On the generalized distance in statistics. Proceedings of National Institute Science, India 2: 49–55

Mosimann JE (1970) Size allometry: size and shape variables with characterizations of the log-normal and generalized gamma distributions. Journal of American. Statistics Association 65: 930–945

Mardia KV, Dryden IL (1989) The statistical analysis of shape data. Biometrika 76: 272–282

Pearce SC (1959) Some recent applications of multivariate analysis to data from fruit trees. Annual Reports East Malling Research Station 1958: 73–76

Pearce SC (1965) Biological statistics. McGraw-Hill, New York, xiii + 212 pp

Pearson K, Morant, GM (1934) The Wilkinson head of Oliver Cromwell and its relation to busts, masks and painted portraits. Biometrika 26: 1–116

Quenouille MH (1952) Associated measurements. Butterworths, London, x + 242 pp

Rao CR (1952) Advanced statistical methods in biometric research. Wiley, New York, 292 pp

Rao CR (1966) Discriminant function between composite hypotheses and related problems. Biometrika 53: 339–345

Reyment RA (1991) Multidimensional palaeobiology. (+ appendix by LF Marcus) Pergamon Press, Oxford, ix, 377 pp

Reyment RA (1996) An idiosyncratic history of early morphometrics. In Marcus LF et al. (Eds.) Advances in morphometrics, NATO ASI series: Series A, life sciences, volume 284, 15–22 pp

Reyment RA (2005) Aspects of applied morphometrics. Zeitschrift der deutschen geologischen Gesellschaft. 154: 263–274

Reyment RA, Banfield C (1976) Growth-free canonical variates applied to fossil foraminifers. Bulletin of the Geological Institutions of the University of Uppsala. N.S. 7: 11–21

Reyment RA, Jöreskog KG (1999) Applied factor analysis in the natural sciences. Cambridge University Press, Cambridge, 371 pp

Reyment, RA, Kennedy WJ (1990) Phenotypic plasticity in a Cretaceous ammonite analyzed by multivariate statistical methods. Evolutionary Biology 25: 411–426

Reyment RA, McKenzie KG (1993) Quantitative genetics in Paleontology: evolution in Tertiary Ostracoda. International Association for Mathematical Geology: studies in Mathematical Geology, Computers in Geology – 25 years of progress. In Davis JC, Herzfeld U (Eds.) University Press, Oxford, No. 13, 155–168 pp

Reyment RA, Blackith RE, Campbell NA (1984) Multivariate morphometrics, Second Edition, Academic Press, London, vii, 233 pp

Robins G (1994) Proportion and style in ancient Egyptian art. University of Texas Press, Austin, 283 pp

Rohlf FJ (1993) Relative warp analysis and an example of its application to mosquito wings. In Marcus LF et al. (Eds.) Contributions to morphometrics, Museo nacional de ciencias naturales, Madrid, 131–159 pp

Rohlf FJ (2002) Geometric morphometrics and phylogeny. In MacLeod N, Forey PL (Eds.) Morphology, shape and phylogeny, Systematics association special volume series 64, Taylor and Francis, London, 175–193 pp

Siegel AF, Benson RH (1982) A robust comparison of biological shapes. Biometrics, 38: 341–350

Sprent P (1972) The mathematics of size and shape. Biometrics, 28: 23–28

Swiderski DL, Zelditch ML, Fink WL (2002) Comparability, morphometrics and phylogenetic systematics. In MacLeod N, Forey PL (Eds.) Morphology, shape and phylogeny, Systematics special volume series 64, Taylor and Francis, London, 67–99 pp

Teissier G (1938) Un essai d'analyse factorielle. Les variants sexuels de *Maia squinada.* Biotypologie, 7: 73–96

Thompson DW (1942 [1917]) On growth and form. University Press, Cambridge, second edition, 1116 pp

Zahn CT, Roskies RZ (1972) Fourier descriptors for plane closed curves. IEEE Transactions on Computers, C-21: 269–281

http://www.springer.com/978-3-540-95852-9

Morphometrics for Nonmorphometricians
(Ed.)A.M.T. Elewa
2010, XI, 367 p., Hardcover
ISBN: 978-3-540-95852-9