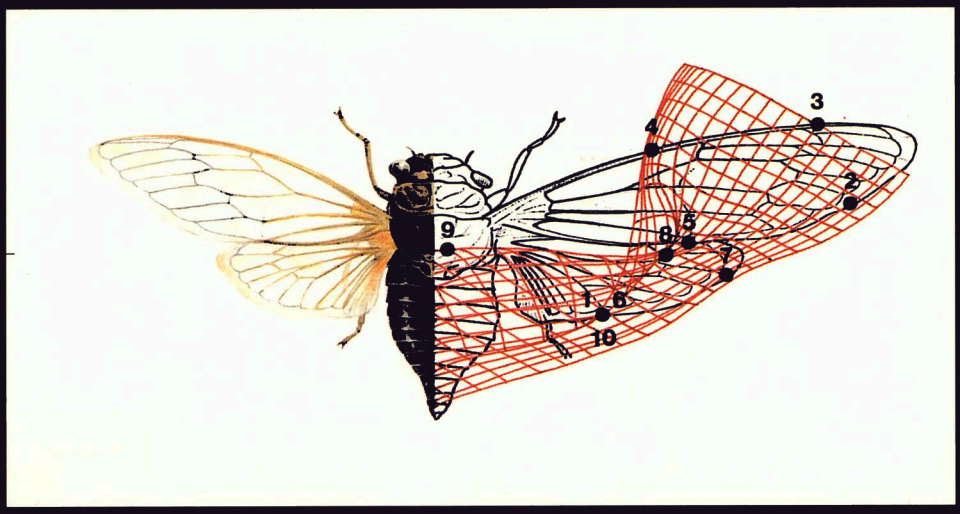
Edited by
Leslie F. Marcus, Elisa Bello
& Antonio García-Valdecasas



CONTRIBUTIONS TO MORPHOMETRICS

MONOGRAFIAS

MUSEO NACIONAL DE CIENCIAS NATURALES
Consejo Superior de Investigaciones Científicas

CONTRIBUTIONS TO MORPHOMETRICS

Edited by

Leslie F. Marcus

American Museum of Natural History, New York

Elisa Bello

Museo Nacional de Ciencias Naturales, Madrid

Antonio García-Valdecasas

Museo Nacional de Ciencias Naturales, Madrid

MONOGRAFIAS

MUSEO NACIONAL DE CIENCIAS NATURALES

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS

ISBN: 84-00-07353-3
Depósito Legal: M-11292-1993
Impreso en España. Printed in Spain
GRAFICAS MAR-CAR, S. A.
Ulises, 95 - 28043 MADRID

INDEX

Page

Part Four: APPLICATIONS

APPENDICES

INTRODUCTION

Morphometry is undergoing major changes on at least two fronts. One is the development of new methods by statisticians working on biological problems, and the other in the kinds of biological questions to which they are applied, especially a new look at development, evolution and variation in the form of organims. Form includes both size and shape. Up to now, size has been dealt with in a satisfactory quantitative manner, but shape has been reduced to a comparison of sizes of parts of organisms. Now, it is possible to dissect shape into linear and non-linear components for homologous landmarks. A number of workshops in the last few years and additional ones being organized are disseminating the results of this effort, and attest to the vigor and interest in morphometrics. We believe it is not too early to describe the emerging changes as radical departures from the past. A new paradigm is now being formulated. We also believe that the store of available problems and methods qualifies morphometrics as a discipline in its own right.

A central theme in the new Morphometrics is to consider landmarks directly, rather than to derive distances from them. The results are reported in organism space rather than in the abstract vector space of classical multivariate statistics and traditional morphometrics. However, classical multivariate techniques are still relevant in testing and inference. The benefits of such an integration are still not as widely appreciated as we would like, but further efforts such as this one will hopefully make the methodology more available. Fred Bookstein has pointed out that we have a "desperate need of a book-length primer" on the new morphometrics.

Much of the literature and examples in systematics are in two dimensions rather than in the three that we observe and study. It is clear from the description of the methods that the mathematics has been developed, algorithms are available, and numerical results are easily obtained for three dimensional data. Three dimensional (3-D) data acquisition is still expensive for the systematist, who now almost routinely uses 2-D video capture systems.

Three dimensional software is being developed and the speed of low cost desk top computers is increasing so that the dynamic graphics required will soon become widely available.

This volume reports on some of the methods and applications available in the new morphometrics. Bookstein in the first chapter has provided us with a good history of methods tracing the roots of the new perspective. Key phrases are "alternative visualizations", and "configuration of landmarks".

Roth in her essay outlines practical necessities in collecting adequate data for morphometric analysis, emphasizing care in photography and measurement. She discusses the relations between our customary 2-D perspectives and the 3-D reality of objects.

In fact there has been a 2-D emphasis in the new morphometrics applied to systematic work to date, and the gap that exists between the availability of data acquisition techniques and expensive computer displays required for the newer kinds of data used in medicine for example, will be closed in the near future.

Becerra, Bello and Valdecasas offer some practical advice on selection of equipment for 2-D image capture for morphometrics based on their experience with a selection of hardware and software available to them. They report also some experiments to give an idea of resolution and repeatability available on the lower cost systems built around IBM PC's and clones, and data acquistion software. We hope that gatherings in the near future will be able to discuss the pros and cons of various 3-D systems as clearly.

Slice offers a contribution dealing with outline data. He gives a useful critique of the use of fractal dimensions in describing two dimensional shapes - using leaves as examples. He offers many practical points in the use of this methodology, and discusses pitfalls to avoid.

Marcus gives some practical applications of classical multivariate statistics in his article, and discusses some relations between univariate and multivariate inferential statistics. The biplot method is explained, and heuristics on relations between student's t and Hotelling's T^2 as well as analysis of variance (ANOVA), and multivariate analysis of variance (MANOVA) are provided. Programs written for the software package MATLAB are included to support some of these methods.

Rohlf has contributed a clear algorithmic development of the technique of relative warps. His article is at the same time a user guide to his Thin Plate Spline Relative Warp (TPSRW) program provided with this volume. He explains all of the steps and interpretations with his now familiar data set on mosquito wings. Practical choices in terms of ontogenetic and exploratory studies are discussed as well. The relative warp technique in his hands is seen as an another way of operating with linear functions of the data, and in this way broadens our view of this family of techniques.

The most important test for a recipe is the tasting, and both Walker, and Loy et al. have done just that. Walker has extended resistant fit techniques to an exploration of landmark allometry and provided useful graphics. He has also contrasted alternative registration methods in both interpretation and presentation of his results.

Loy *et al.* apply Bookstein shape coordinates in their analysis of systematic diversity of European moles. They are interested in phylogenetic reconstruction, sexual dimorphism, and intra-specific variation in the skulls of these highly specialized mammals. This application also serves to illustrate the use of classical multivariate statistics with data derived from the new morphometrics.

Becerra has provided a useful discussion of electronic mail and other communication possibilities over BITNET and the Internet. His article serves as a primer for those new to these topics.

Finally the latest versions of GRF, TPS and TPSRW by Rohlf are provided on a disk included with this book. A completely new program TPSREGR is provided by Rohlf. See the Appendix and README file with that program for a discussion of its features. The latter especially, is a considerable revision of the original accompanying the Michigan Morphometrics Workshop volume. An appendix explains their installation and use. Also see the Appendix on how to obtain newer up to date versions.

Some programs are provided in MATLAB which means that they can be run on IBM PC´s or clones, Mac´s, work stations and other platforms if one has the package MATLAB avaible. Programs are included to produce the Biplots in the paper by Marcus, and a program TPSNEW and TPSRWZ3 that do thin plate splines and relative warp computations following the output and steps very closely in Rohlf's article. Other software included are documented in the README file included on the diskette accompanying this volume.

LESLIE F. MARCUS
ELISA BELLO
ANTONIO GARCIA-VALDECASAS

New York, Madrid
December 1992

PART ONE

HISTORY, CONCEPTS, DISCUSSION AND CRITICISM

A BRIEF HISTORY OF THE MORPHOMETRIC SYNTHESIS

FRED L. BOOKSTEIN

Center for Human Growth
University of Michigan
Ann Arbor, MI 48109
FRED L. BOOKSTEIN@UM.CC.UMICH.EDU

CONTENTS

ABSTRACT

The modern morphometrics of landmark data represents a surprisingly recent synthesis of two originally divergent methodological styles. One contributory stream is the tradition of multivariate biometrics originated by Francis Galton, developed further by Karl Pearson, and brought into its current form by Sewall Wright. These approaches emphasize the geometry of the covariance matrix over either the gometric organization of the measures or their biological rationale. The other stream, usually attributed to D'Arcy Thompson but actually dating from Renaissance art, emphasizes the direct visualization of changes in biological form; until quite recently it lacked a statistical method. The two approaches lead to quite different versions of morphometrics because they tap quite different notions of homology: at root, they represent incompatible channels of data. Earlier approaches to a biometrics of organic form applied the multivariate metaphor rather arbitrarily to various subsets of the available information.

But the goal of combining the two forms of biometric modeling arose – especially among amateurs – with remarkable regularity throughout the century; finally, from the late 1970's through the 1980's, the two families were firmly fused. The key strategic decision was the restriction of the data base to locations of discrete landmark points that sampled transformations at the same time that they sampled individual forms. A combination of the geometry of the mean landmark shape with the geometry of the covariance matrix of these shapes leads to the quantification of transformations in tractable form and to the visualization of most conventional multivariate maneuvers as transformations. The synthesis, carried out by Bookstein, Kendall, Goodall, Mardia, and others, carefully combines ideas and mathematical tricks from statistical "shape space", multivariate analysis, algebraic geometry, and interpolation theory. This paper summarizes the separate histories of the two morphometric traditions, the salient features of the synthesis, and the lessons of this history for the larger context of methodological advances in biometry.

This essay is a slightly modified version of one to appear in the commemorative Volume 100 of Springer´s "Green Series", *Lecture Notes in Biomethematics*, edited by Simon A. Levin.

INTRODUCTION

For most of the twentieth century, techniques for the biometric analysis of organic form fell into one of two incompatible styles. In the first, more indigenous style, a direct extension of techniques introduced into statistics by Galton, Pearson, and their heirs, conventional multivariate techniques were applied to a diverse roster of measures of single forms. The only algebraic structures involved were those of multivariate statistics, limited mainly to covariance matrices; no aspect of the geometric organization of the measures, or their biological rationale, was reflected in the method. Analyses of this mode led at best to path diagrams, not to sketches of typical organisms expressing the developmental or functional import of the coefficients computed.

In the other class of shape analyses, often associated with the name of D'Arcy Thompson but actually dating from the Renaissance, changes of biological form were visualized directly as distortions of Cartesian coordinate systems that accorded with a pre-assigned biological homology. Such analyses were inextricably graphical; several generations of bricoleurs failed to provide a corresponding statistical method. Whereas in the first approach homology pertains to the values extracted by ruler, in the second it refers to the pairing of "corresponding" locations of bits of tissue. The incompatibility between these two main styles of quantification derives ultimately from this discrepancy between fundamental metaphors for what is being measured.

Recently these two broad families of techniques have been fused in a surprisingly brief and peaceful methodological development. The key to the synthesis was the restriction of the information being analyzed to the locations of discrete points, landmarks, that bore Cartesian coordinates but that also were declared to be biologically homologous from form to form of a series. Over the decade from the late 1970's to the late 1980's, Thompson's transformation grids, as applied to landmark configurations, were quantified in a statistically tractable form. The statistical analysis of landmark locations was shown to be expressible in geometric diagrams directly interpretable in the original picture plane, and the results of the statistical analysis were made commensurate with the analytics of the deformation analysis with the aid of an ancillary quadratic form encoding the sample average positions of the entire landmark configuration.

The resulting morphometric synthesis is of full statistical efficiency, permits explicit tests of many biologically interesting features, and supplies statistical

equivalents to, or statistical instructions for, the great variety of graphical techniques that had been previously developed by amateurs. This essay briefly recounts the history of algebraic and geometric manipulations that culminated in the current state of morphometrics. My subject is the accumulation of insights into the logic of measurement: the history of a method, not of findings.

BIOMETRIC ANALYSES OF SIZE AND SHAPE MEASURES

Of the two reviews of independent developments just sketched, the more ironic is the easier to write. Modern biometrics is a grand intellectual structure, with applications from population genetics through psychology and into the social sciences. But whereas it arose in response to specific tasks of size and shape analysis, the most successful of its techniques are incapable of making any use of such geometric origins for the data. The power of biometric methods for broader applications – the fact that discriminant function analysis, for instance, works as well in psychiatry as in botany – owes to its discarding half the information of the biometric context, the information that is peculiarly *bio*metric, at the outset. This missing information will not be restored until quite nearly the end of our history.

Throughout the early history of today's biostatistical methods, data for exemplary demonstrations typically derived from biological size measures. The original quantitative study of development, for instance, was de Montbeillard's (1760) tabulation of the height of his son (see Boyd, 1980). Quetelet's uncovering of the normal distribution in a social context relied on measures of height, weight, and the like. And, of course, Francis Galton's original demonstration of regression used the heights of 928 children and their parents. Following Duncan (1984), I would suspect all this owes to the origin of these thrusts in the need for "social measurement" long before the idea of biometric statistics could be formulated. Generals and tailors needed to understand human size variability millenia before quantitative biology was more than an eccentric hobby.

The independence of multivariate algebra from the biometric context in which it originated was noted very early on in the development of multivariate statistics. Recall that Galton emphasized two related but distinct aspects of the relation between parental and child height: the fact of regression (that is, the true, linear causation of what we now call the "systematic part" of child's height by midparent height) and the convenience of a summary statistic of "co-relation". Well before the turn of the century, Edgeworth and then Karl Pearson and his colleague W. F. R. Weldon were emphasizing the usefulness of this second formulation for pairs of biometric variables, such as alternate size measures of the same organism, for which the true causal model was not at all that of Galton's hereditarian exemplar. Still, the principal ideological concern of this school of thought, usually

identified with the political thrust of the eugenic movement (Mackenzie, 1981), was the effort to restrict regressions to a context of true causation. This purpose was distinctly oblique to G. Udny Yule's (1895) construction in which regression analysis consisted simply of "fitting a plane to the data" for purposes of easing prediction at the expense of explanation. Matters were not helped when Ronald Fisher's algebra of *explained variance*, a terminology suited to the context of agricultural experimentation, proved to apply to the decomposition of sums of squares underlying multiple regression, which, properly construed, "explains" nothing – but this theme is oblique to our main story.

The meaning of regression and correlation in biometric studies of size and shape measures was obscure until the 1920's, when Sewall Wright applied his considerable analytic and intellectual skills to their clarification. His method of *path analysis* (see Wright, 1968) was developed to unify studies of inheritance of quantitative characters and of the correlations among simultaneously measured suites of characters. In the shared formalism, observed correlations were the algebraic composite of patterns of mutual determination of data by observed or unobserved *factors*. Wright's conception of the role of correlations in biometrics is still, in my view, the only coherent approach to their application in the biological sciences (see Bookstein *et al.*, 1985, or Bookstein, 1991).

While Wright was developing his strictly causal models, and enlarging their range to include selective forces and genetic drift and diffusion, the opposing tradition (regression as least-squares prediction) was not dormant. Of the many intellectual developments which branched from the "general linear model" (multiple regression, analysis of variance, and their common generalizations), several are crucial to modern applied statistical practice, including econometrics, response-surface analysis, and psychometric factor analysis. While none of these have recurved to enrich biometrics in any central way, at the same time yet another development was arising in the context of morphometric data. The technique of *discriminatory analysis* originated in Fisher's classic data set of four size measures of *Iris* flowers. In phrasing his problem of "discrimination" as the maximization of a certain variance-ratio, at root a ratio of statistical likelihoods, Fisher failed to notice that he was once again denying the origin of the biometric task in any coherent causal model. This was confirmed shortly afterwards when Harold Hotelling (1936) showed how discriminant function analysis was a special case of *canonical correlations analysis*, a technique that had arisen in the context of psychometric statistics to make sense of group differences in "profiles" on any outcome whatever, regardless of the style of measurement and regardless of the nature of the true factors, if any, controlling the phenomenon under study.

By the 1960's, then, the discipline of biometrics found itself in a context of considerable internal contradiction. The core collection of techniques – regression, true factor analysis, discriminant function analysis – had arisen in the context of a strictly morphometrical question, yet in their current algebraic unfolding *there*

was no role for any geometrical information at all. The question of whether the algebra of covariance matrices and design matrices did justice to the biological hypotheses so investigated could not be posed.

The dilemma is presented quite neatly, if inadvertently, in the first pair of publications known to me that actually claimed to be about "morphometrics": R. E. Blackith's (1965) essay of that title and his 1971 book *Multivariate Morphometrics* with Richard Reyment. In both these texts, morphometrics is mainly the interpretation of matrix manipulations in vaguely functional biological terms. Summarizing the field as it had ramified over the preceding half-century, these authors were quite free to ignore the origin of the variables under study. The nature of the measures – lengths, angles, titres, proportions, whatever, in any combination – made no difference for the matrix mechanics: all were thrown into the same vortex of canonical analyses and clusterings. Thus there could arise no discipline for the formulation of those variables. In a related literature, the applied field which supplied morphometric data to the greatest accuracy, craniometrics (along with its alternative incarnations anthropometrics and cephalometrics), seems never to have considered what might be a reasonable approach to their provenance. Distances, angles, ratios, areas – all are combined helter–skelter in unitary matrix analyses from which biological insight is presumed to emerge by inspection of tabular results or ordinations.

The clumsiness with which the methods of this suite apply to the actual data of size and shape in which they had been conceived did not go wholly unnoticed. Rather, from mid-century on, several thoughtful biometricians attempted to modify the dominant matrix methods so that when interpretations in terms of size and shape were possible they might be called to the scientist's attention without any more distortion than was absolutely necessary. Jolicoeur (1963), Hopkins (1966), Burnaby (1966), Mosimann (1970), and others investigated the interactions of the biologist's intention with matrix operations as applied to true measures of size and shape. For instance, the (true, causal) phenomenon of *allometry*, dependence of shape on size, can (sometimes) be detected in variation of the coefficients of the first principal component of logarithms of size measures; analysis of "shape" can proceed (under fairly stringent conditions, and with limited power) using vectors of ratios of size measures; analysis of shape in a different sense, now no longer size-independent, can proceed by referring to residuals of the raw data from their allometric regressions; and so on. This literature is summarized and assorted in Bookstein *et al.* (1985), and its semantics is dissected in Bookstein (1989b).

Still, by about 1980 the inescapable mismatch was clear to many of us between the matrix operations of the dominant tradition, however modified for "size and shape" work, and the very reasonable sorts of questions about morphometric phenomena that had been asked of the raw data all along. In a phenomenon typical of such periods of professional stress, new techniques began to spring

up only to be found deeply flawed, or otherwise misadvertised, shortly afterward. Among such techniques were (alas) my idea of "shearing" (size-free shape discrimination from rotated principal component analyses, Humphries et al., 1981), the vain hope that Fourier and other orthogonal functional analyses of form might result in "characters" (cf. Rohlf, 1986), the multivariate analysis of Cartesian coordinate data without any preparation (Corruccini, 1981), and several others. A review article of the time (Oxnard, 1978) summarized morphometrics as a grab-bag of techniques borrowed from a great variety of sources – statistics, engineering, optics, psychometrics – without any coherence of its own. For any such coherence to arise, the field would have to be rebuilt from first principles emphasizing the origins of the data (quantitative observations of actual biological form) as much as the algebraic machinery of its statistical analysis. But where to begin? – what questions should be placed at the foundations of morphometrics, to set the rules of discourse prior to particular applications?

THE STUDY OF SHAPE TRANSFORMATION

The coherence lacking in the morphometrics of my graduate years was born, though not without forceps, out of a completely different tradition than the biometric: the systematic contemplation of biological shape *change* as a phenomenon in its own right. While this idea is usually associated with the famous treatise *On Growth and Form* (1917) by the British naturalist D'Arcy Thompson, it is actually hundreds of years older than that. The first "transformation grids" reflect efforts of Renaissance artists to comprehend the variability of the human forms that they were just beginning to reproduce realistically. Figure 1, for instance, from Albrecht Dürer's *Vier Bücher von Menschlicher Proportion* of 1524, demonstrates a surprisingly broad exploration of diverse types of "transformation grid," both affine and localizable, in the effort to explore the limits of normal variation and the strategies of effective caricature.

This formal theme, shape transformation as the explicit object of biometric discussion, was first clearly set forth in the famous Chapter XVII of Thompson (1917), *On the Theory of Transformations, or the Comparison of Related Forms.* Thompson's goal is a distinctly Victorian one, perhaps too Platonic for the modern taste:

> [If] diverse and dissimilar [organisms] can be referred as a whole to identical functions of very different co-ordinate systems, this fact will of itself constitute a proof that variation has proceeded on definite and orderly lines, that a comprehensive 'law of growth' has pervaded the whole structure in its integrity, and that some more or less simple and recognisable system of forces has been in control...

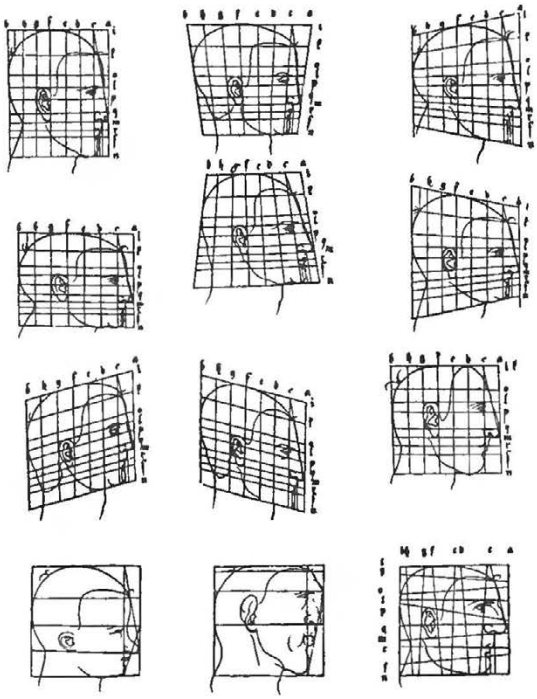


Fig. 1
Pre-Cartesian transformations. From Albrecht Dürer, Vier Bücher von Menschlicher Proportion, 1524.

Indeed, the figures which he himself published show a clear dominance of the Platonic thrust of homogeneity over accuracy or even realism in the representation of actual data. Thompson's hope that these figures would help unveil the origins of form in force was never realized, and while several later generations of quantitative biologists were tempted by this graphical style, it proved never to lead to quantification in the global mode that Thompson had intended. For a historical review of the "vicissitudes" of this method since Thompson's publication, see Chapter 5 of Bookstein (1978).

From the vantage point of 1992, it is possible to characterize the assortment of earlier attempts at a proper biometrics of transformation by the nature of the compromises they made. We shall see below that the morphometric synthesis involves many separate themes in the biometrics of shape: representation of variation of shape and size as well as mean effects, coverage of a full range of potential shape descriptors in an even and "unbiased" fashion, and production of distinctive features of such changes or variation in multiple diagrammatic forms permitting their separate viewing and also their arbitrary combination in composite processes. All this needs to be under the control of the same conceptual unity of descriptions that is presumed the case for ordinary variables: all comparisons must be "of like with like". The innovation of the 1980's consisted in a single formalism allowing

of all these alternative emphases; it is no criticism of those who came before that they had not stumbled upon the appropriate statistical geometry.

Sneath & Sokal (1963), for instance, presented realistically drawn Cartesian transformations between holotypes, but argued (following Medawar) that such visualizations did not lead to "features" or to measures of "distance", and so turned elsewhere for the multivariate distance measures that were supposed to lead to taxonomically appropriate ordinations. A few years later, Sneath (1967) attempted to convert smoothed models for these grids into a trend-surface-based distance function; but there was no possibility of interpreting the resulting coefficients in geometric terms. Huxley's (1932) method of "growth-gradients" would occasionally lead to suggestive Cartesian transformation diagrams, but begged the question of an appropriate coordinate system. Bookstein's method of biorthogonal grids (1978) provided shape comparisons in a canonical coordinate system but was not consistent with visualizations of "standard error" or any other notion of sampling variance for the features so displayed. Oxnard's method of displaying single principal components of multivariate size measures as grids (1973) represented statistically reliable shape features in diagrams whose verbalization (e.g., "cranio-lateral twist") is obscure: what family of descriptions are we drawing descriptive phrases like these from, and how much of that "twist" do we have? Yet other methods, such as Lohmann's "eigenshapes" (1983), which could be thought of as transformations of the boundary of a form, failed to accord with prior knowledge of biological homology, but instead construed it in an operational fashion that, however effective for ordination or correlation with ecophenotypy, nevertheless did not permit interpretation in biological terms.

The earliest applications of tensor analysis in morphometrics, such as that of Richards & Kavanagh (1943), while strongly suggesting developmental interpretations, did not permit group-level operations such as averaging or assessments of variation; and the later finite-element methods, such as that of Lewis *et al.* (1980) or Bookstein (1984a), displayed "features" the provenance of which was an unknown function of the (arbitrary) division into "finite elements" that underlay every set of specific computations. The methods of Procrustes analysis (optimal least-squares superposition of shapes), which were entering applied morphometrics just as the synthesis was being produced on the pure side, produced "features" of one kind only (vectors of displacement of single landmarks) and were inconsistent with the usual sorts of multivariate explanations (for instance, allometric and growth-gradient models).

In hindsight we can see why the morphometrics of the 1950's through the early 1980's was so confused. There was no agreement about what constituted an appropriate *analysis* because there was no proper theory of what constituted *the data*. Oxnard's (1978) review article, for instance, which dealt with data in the form of images, had virtually nothing in common with the approach of Blackith & Reyment (1971), which treated data in the form of geometric variables measured by ruler, planimeter,

or protractor; and my first publications of the "method of biorthogonal grids" in the late 1970's were unaware that the statistical problem had to do with the representation of the raw data (in this case, whole configurations of landmarks) in a space whose dimensions would *be* transformations, not with the depiction of single changes *as* transformations. My preliminary statistical method for triangles (Bookstein, 1982a,b), lacking only the corresponding distribution theory, never referred to vectors of variables, nor did it hint at any appropriate extension even to pairs of triangles, let alone to landmarks considered without lines connecting them.

In short, none of us realized that the multivariate tradition could not apply properly (i.e., canonically, with full efficiency) to landmark data until a canonical way were found to make whole landmark configurations into "variables", and none of us thought to pursue the analysis common to alternate visualizations rather than the argument that some visualizations were "better" than others. When analyses appeared to work in particular examples, we could not state what it was that caused us to trust in them, nor could we assure ourselves that other analyses, just as cogent, would result in similar findings. In comparing methods for analysis of outlines to methods for analysis of landmark data, no-one was able to say where lay the essence of the difference. (We now know that the essential feature is the finite-dimensionality of the complete description of a landmark configuration.) Thus, a whole collection of earnest workers, some amateurs, some professionals, circled around the solution that was to come, without ever realizing the crux of our collective problem.

THE MORPHOMETRIC SYNTHESIS 1983-1989

Suddenly, without any premonitory ferment, the earlier biometric barriers were circumvented by the combination of many earlier methods in new ways. The breakthrough began, as statistical breakthroughs often do, when it was realized what constituted the appropriate "simplest case": not a short list of distance measures, but instead the simplest configuration of landmarks – a *triangle*. We knew that statistical analysis of triangles by distances (for instance, the lengths of the edges) was not conducive to visualization of effects on these forms (by strain-crosses, pairs of distances at 90°). Thus several of us were searching at the same time for a better multivariate statistical analysis that would wrestle with the landmark locations directly, rather than in the form of the nonlinearly derived lengths, length-ratios, principal strains, etc. This better synthesis emerged between 1983 and 1989 as an essentially complete framework for the analysis of landmark locations as raw data.

The important contributions during this brief period when the discipline was synthesized include a paper of mine (Bookstein, 1984b) introducing the shape coordinates for triangles and showing how shape differences can be weighed by formal T^2 test; Goodall's 1983 dissertation, deriving the equivalent F-ratio while avoiding any size-standardization; and Kendall's (1984) announcement of the global shape spaces to

which Goodall's and my methods inadvertently applied as statistical metrics in tangent spaces (linearized feature spaces). Our joint publication in the first volume of *Statistical Science* (Bookstein, 1986, with commentary) proudly announced the convergence of all three of these approaches on one single foundation for the morphometrics of landmarks. This core of material has since been formalized further, in a different notation, in Goodall (1991). Meanwhile, one particular interpolation function, the *thin-plate spline* (Bookstein, 1989a), turned out to support a feature space *for* these shapes in an almost miraculous way: A quadratic form embodying the mean landmark configuration served to specify a basis for sensibly decomposing variations around that mean. I am not aware of any serious problems with this synthesis or of any informed attacks upon it. Its most extensive exposition is my monograph of 1991; the *Proceedings of the Michigan Morphometrics Workshop* (Rohlf & Bookstein, 1990) provide a link to the language of systematics. There is a useful chapter-length overview in Reyment (1991). We are all in desperate need of a book-length primer.

As the present essay is an experiment in intellectual history, rather than a medium for explaining how to do modern morphometrics, I shall summarize the actual content of the synthesis only briefly, in this paragraph and the next three, before preparing to show how it sits atop most of the morphometrics that had gone before. The *shape* of a set of K landmarks in a plane can be considered as a point in a well-characterized elliptic manifold of dimension $2K$-4 (cf. Fig. 2 A). In small

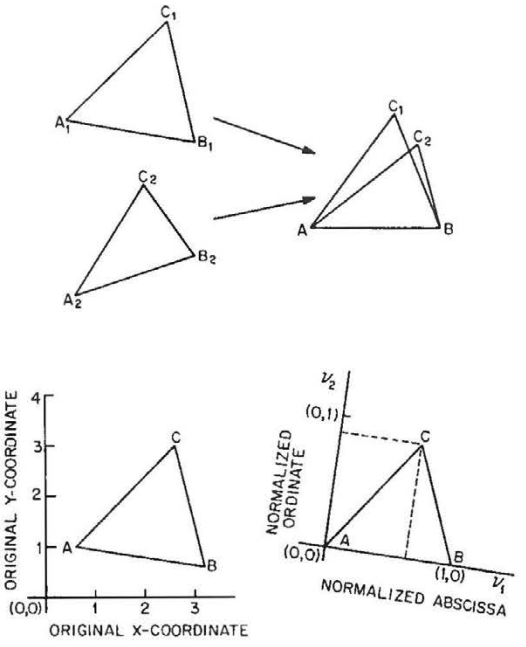


Fig. 2 A

When the positions of two landmarks are fixed, the shape of any triangle is archived by the pair of coordinates of the third landmark. (From Bookstein, 1991, Figure 5.1.2).

regions of this space, ordinary multivariate maneuvers may proceed by the usual machinery of analyses of variance, regressions, discriminations, and the like, as applied to any convenient basis for the tangent space that linearizes "small" shape variations. Under convenient null hypotheses, distributions in this space can be calibrated according to so-called *Procrustes distance*, arc-cosine of the root mean squared distances between the positions of paired landmarks when each configuration is scaled to central second moment unity and when they are rotated and translated to the superposition of least such distance.

But this distance cannot serve effectively as the multivariate "interspecimen distance" beloved of taxonomists in fact, *no* formula for distance can do so; the problem of describing biological shape variation is subtler than that simplistic multivariate model. As matters stand today, there appear to be a minimum of *three* distances involved. For size, log Centroid Size appears to be satisfactory in most applications to landmarks. For biological work in shape space, two distances seem to be required. One, usually log anisotropy, is taken between projections of the two landmark configurations onto the uniform subspace (in directions which vary from algorithm to algorithm; see Bookstein, 1991, Sec. 7.2). The third distance represents position in the complementary subspace of nonuniform transformations, and may correspond to a quadratic form representing some power of the bending-energy matrix. These three distances relate among themselves, through the observed or computed coordinates which reproduce them, by ordinary biometric covariance structures, and they relate to putative exogenous causes and effects by Wright-style path models just as any other quantitative characters would. This point of view is discussed at several places in Bookstein, 1991, from Chapter 1 on.

What makes the synthesis supersede so many of the earlier, partial approaches, even though multivariate "distance" is irreducibly ambiguous, is the existence of a few particularly convenient bases for this space that together support all the visualizations needed for biological interpretations of the formal statistical analyses. The formulation of these bases crucially incorporates the mean landmark configuration. The synthesis is unusual among multivariate methods in this central role of the multivariate mean vector for interpreting variance-covariance matrices around it. In any of these bases, each dimension points in the direction of multiples of one single transformation of a mean form, just as Thompson might have envisioned had he been statistically inclined. Our "features", each of which deforms the mean configuration into some variant, can be depicted (unambiguously) by graphics: some by Cartesian grids, some by simpler vector diagrams. That is, the Cartesian grids are not properties of the data; they are properties of the representation of the data by specific basis vectors. Vectors come first, grids later.

In most of the current implementations (but see Goodall & Mardia, 1991), the basis vectors come in pairs corresponding to the two dimensions of circular symmetry needed to handle their possible application in any direction of the plane.

Particularly convenient (owing to their combinatoric flexibility) may be *shape coordinates*, the shapes (realized as complex numbers) of any K-3 triangles that rigidly triangulate the landmark configuration (cf. Fig. 2A, 2B). For study of large-scale effects on shape, one pair of dimensions is usually reserved for the uniform shears, those which leave parallel lines parallel. Distance in this plane may often be measured usefully by *log anisotropy*, the logarithm of the ratio of axes of the ellipse into which a circle is deformed. The remaining $2K$-6 dimensions can be visualized as K-3 pairs in several ways. One is as a collection of K-3 residuals of point locations after an affine Procrustes fit (Rohlf, 1992). But I prefer the

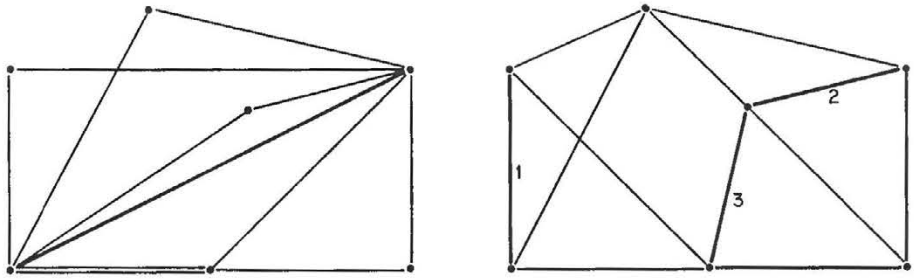


Fig. 2 B

The shape of any configuration of landmarks is archived by the shape coordinates of any set of triangles that rigidly triangulates them. (left) Five pairs to one baseline. (right) Two pairs to baseline 1, two baseline 2, one to baseline 3. The multivariate statistics of shape space can be carried out in a manner independent of all the arbitrary choices involved here beyond the original arbitrary choice, that of the landmarks to be located. (From Bookstein, 1991, Figure 5.2.1).

set of *partial warp scores*. These are 2-vector multiples of eigenfunctions of a particular quadratic form, the *bending energy*, which represents a biological notion of "localizability" in an algebraically convenient form. (The idea of the bending energy was borrowed from the literature of surface interpolation; its eigenanalysis, along with all the rest of the synthesis, rests on indigenously biometric methods.) Whatever the basis chosen, multivariate analysis proceeds by the usual multivariate matrix methods; but then all findings are diagrammed back in the plane of the data by graphics of displacement or deformation each corresponding to one of the directions of shape space in the vicinity of the mean form. For an example, the decomposition of a group mean shape difference by its partial warps, see Fig. 3.

The consensus involving these new methods is important to note because of three shared formal properties that, collectively, obviate most of the arbitrariness that had bedeviled earlier approaches to the same data. The crucial features of the morphometric core are *efficiency*, *complete coverage of feature subspaces*, and *directional symmetry of embedded distributions*.

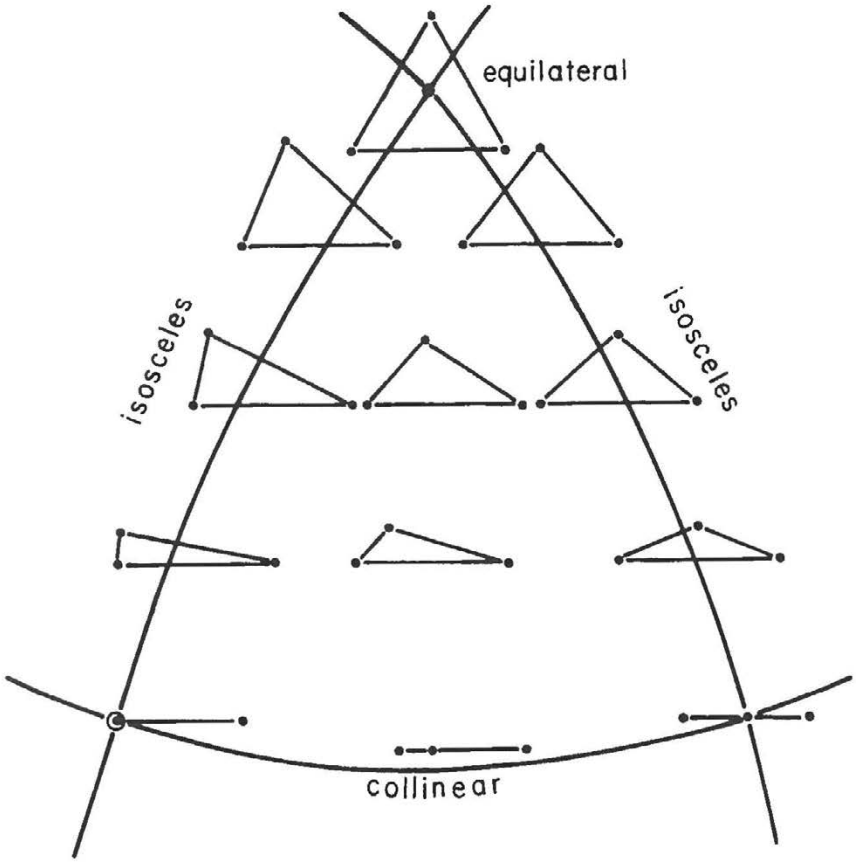


Fig. 3
The relation of Kendall's "spherical blackboard" to the shape coordinates. In the large, shape space is curved. (From Bookstein, 1991, Figure 5.6.1)

Efficiency. One corner of this common foundation is the demonstration by elementary theorem, in my 1986 paper, that the "shape space" common to these schools incorporates the linearized multivariate statistics of all possible "traditional" shape measurements of the same landmark locations. This guarantee of efficiency "in all directions," "in all linearizable features" is perhaps the most important practical consequence of the methodological consensus. Methods have been created for translating these equivalent techniques from one notation to another and for detecting the ways, often subtle, in which other sets of variables *lose* information or efficiency with respect to these optima. For instance, the older studies of single distances and distance-ratios now come under the purview of theorems explaining in advance how their efficiencies may be computed as functions of the mean form. For planar data, the extremes of strain-ratio (ratio of

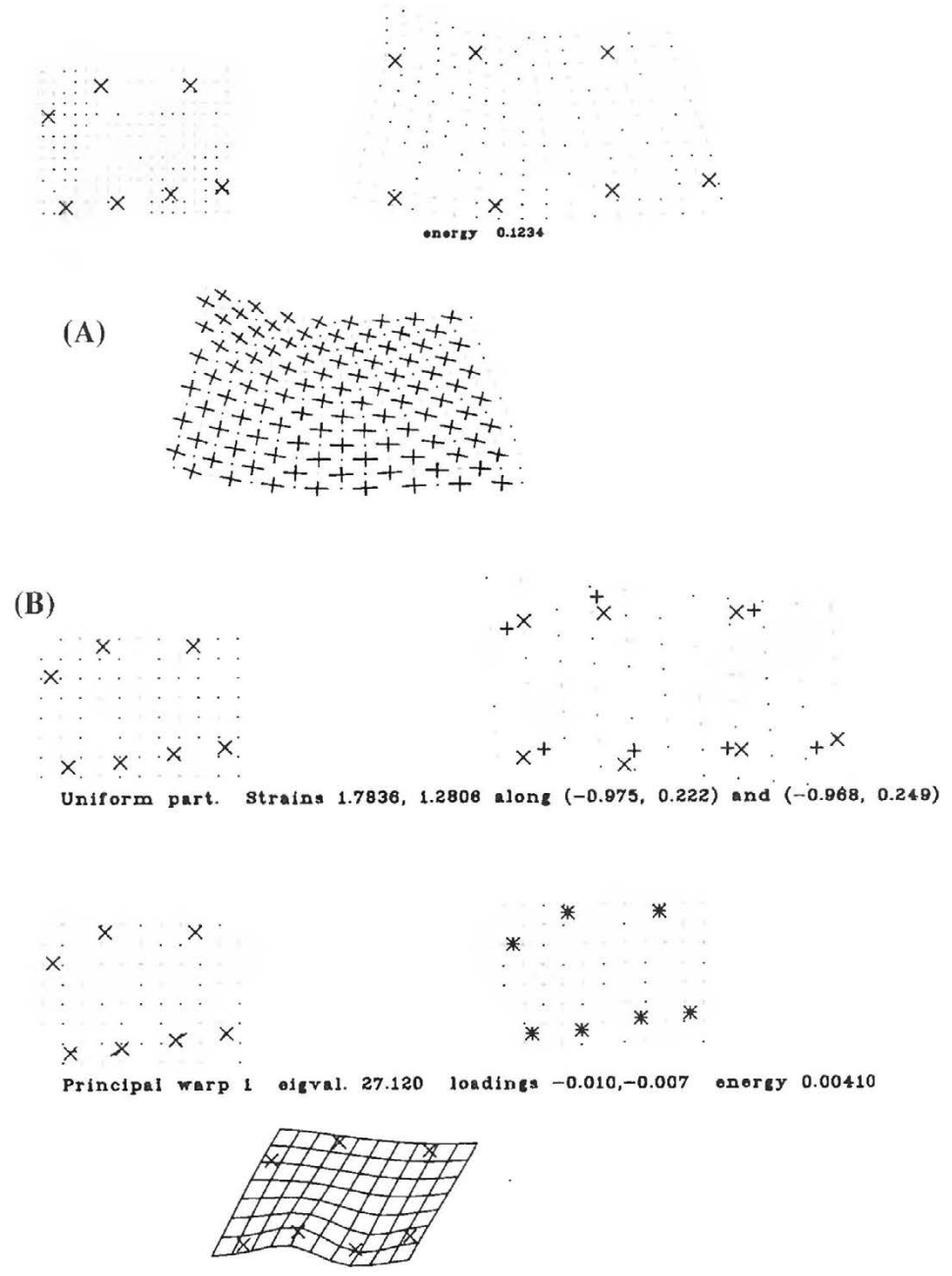


Fig. 4

(A) Thin-plate spline interpolant for mean neurocranial growth in 21 male laboratory rats from 7 to 150 days of age. Below, the asociated tensor field, displayed by way of its principal axes.

(B, continued)

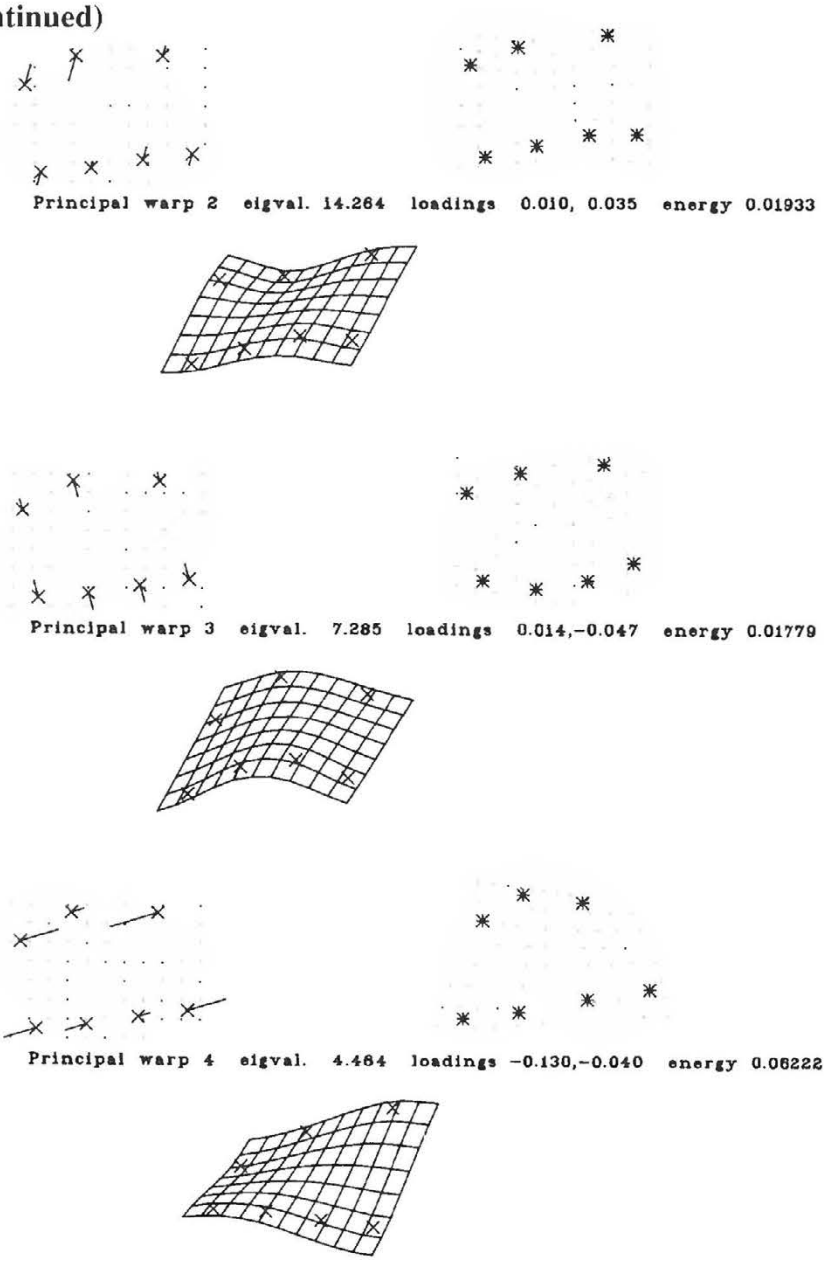


Fig. 4 (continued)

Change in the ten dimensions of shape space for this set of seven landmarks, displayed as five pairs of two dimensions (the uniform component and all four partial warps). Below each partial warp, the underlying principal warp, shown as an actual thin plate. Landmarks (from lateral cephalograms), clockwise from lower left: Basion, Interparietal suture, Lambda, Bregma, Sphenoethmoid synchondrosis, Intersphenoidal suture, Sphenooccipital synchondrosis. (Data from Bookstein, 1991, Appendix 4.5, omitting landmark Opisthion).

corresponding distances between weighted averages of landmarks) from one mean landmark configuration to another may always be expressed along *transects of triangles*, distances measured from single landmarks to the weighted average location of another pair. The net T^2 statistic for a test of group mean shape difference, however, does not reduce to the comparison of two suitably selected exemplars of these transects, one of largest ratio, one of smallest. That net T^2 instead pertains to an appropriately rank-reduced multivariate test for a vector of ratios among themselves of *all* edge-lengths within the configuration. Note, too, that the extreme ratios are typically absent from the basis for the vector space supporting the T^2. It is the joint distribution of the edge-ratios that supports the correct statistical interpretation, not their comparisons separately.

Complete coverage of shape space. When a shape change is detected by this single T^2 statistic, or the equivalent F-ratio (Goodall, 1991) or likelihood (Mardia and Dryden, 1989), the morphometrician's task becomes the specification of features by which this change may be tied to biological explanation. As effects upon shape are of indefinitely wide variety, so, too, are the types of features to be inspected should the hypothesis of isometry (no shape difference) be rejected. Many of these emerged years or decades ago as isolated methods all their own; the advantage of the synthesis is their joint expression as multiple estimations, often statistically nested, in a common format. For configurations of K landmarks, the variants of current interest include the rigid motion of certain subsets of the landmarks with respect to the remainder (each such possibility is associated with a descriptor lying in a subspace having dimension 3 for planar data, and there are 2^{K-1}-K-1 such subspaces); the individual displacements of similar lists of landmarks (having the same number of subspaces, each now of dimension twice the count of "moving" points); the two-dimensional subspace of uniform shears; the eight-dimensional subspace of quadratic growth-gradients; and the scale-specific features of nonlinearity themselves linearized in the principal warps (those "localizable" eigenvectors of bending energy). These are all phrased as group comparisons or exogenous covariances. There are also methods available for studying intragroup morphometric covariances, including single-triangle or uniform factor models and models for within-group features of localized variation. A particularly interesting factor, either exogenous or endogenous depending on the space in which the analyst is working, is size. The associated multivariate methods include techniques for handling the usual questions about allometry, and even dictate a preferred size measure ("preferred" according to a plausible null model of shape noise) corresponding to Mosimann's (1970) theorem restricting size-shape independence. To repeat the advertisement, all these are exposited in Bookstein, 1991.

Directional symmetry. In all these applications there is a geometry of features inherited from the geometry of landmarks. All the methods of the synthesis are circularly symmetric in their weighing of directions in shape space and in all the natural subspaces. By this I do not mean that the data must be modeled as somehow

circularly *distributed* – although such models are available to be tested or rejected – but that the methods explore all directions of variation in shape space using the same metric that applies to the construction of shape space per se, prior to consideration of any covariances. Kendall (1984) shows how this metric is an embedding of the natural Euclidean metric that would be applied to the landmark locations prior to their reduction to equivalence classes of shapes. On a background of n-1 landmarks unchanging in position, the variation of "shape" in the standard construction of shape space is circular (isometric with Kendall's metric) whenever the variation of the real location of that nth landmark is circular in its own picture plane. In other words, in the standard construction the geometry of shape space does not distort displacements of single landmarks as a function of direction.

Beyond this geometry of single point-displacements, there are many other subordinate geometries that may be nested as hypotheses under the standard shape-space construction; methods for these more specialized applications must be circularly symmetric as well. Uniform shears of equal extent, for example, result in displacements to an equivalent distance (as measured by log anisotropy) in the appropriate invariant subspace of shape space, regardless of the principal directions of the shear. In the limit of small changes, these are exact Euclidean circles about the identity transformation in that subspace. Thus all such changes are detected by the common T^2 or F with the same full efficiency, regardless of direction. The same is true of single landmark displacements used to model deviations from the more simplistic Procrustes models.

LESSONS FROM HISTORY

When I began my work in morphometrics, with the book-length attempt at a new foundation, *The Measurement of Biological Shape and Shape Change* (Bookstein, 1978), no-one understood the difference between the problem I was attempting to solve and the problem I should have been solving instead. The entirety of the new methods included in that volume would be discarded in the course of the synthesis that now stands so sturdily on its own.

Of the critique presented in that earlier exposition, nearly all has proved insightful; of the solutions proferred there (the constrained principal components for analyzing closed outlines, the method of biorthogonal grids), nothing much remains. The problem blocking derivation of a landmark-based morphometrics was not, as I erroneously (if understandably) claimed in 1978, the construction of a canonical coordinate system for D'Arcy Thompson's grids. The *real* problem was to construct a statistical space for the finite-dimensional manifold of the landmark configurations themselves such that each vector connecting two points in that space corresponded to a unique diagram of a deformation. Once directions in that space could be named and their statistical reifications assessed, pictorial representations of actual effects on real shapes would follow.

The problem, in other words, was properly to demarcate the biological from the statistical aspects of morphometrics; but for that act of intellectual geodesy it was inappropriate to follow Thompson's lead. In suggesting an elegant graphic for the biologist's intuition of shape *change*, Thompson ignored the problem of shape *description*. Conversely, in suggesting elegant techniques for the manipulation of shape descriptions, the multivariate school ignored the problem of making biological sense of shape differences. Once the embryo of a connection between these two problems could implant itself in the biometrical literature – once we realized that we needed a general descriptive system for general statistical contrasts in the shape space of general landmark configurations– the solution could be written down almost as quickly as we could assemble data sets to serve as examples. Throughout the 1980's, every data set to which the synthesis was applied – plants, rat skull growth, orthodontics, cardiac contraction, brain images – led to findings representing clear methodological advances in all the fields sharing a concern for the crucial methodological lacunae.

There are lessons in this synthesis, then, both for biometrics per se and for methodology more generally. The lesson for biometrics is perhaps the simpler: let the explanations at which the biologist ultimately needs to arrive (in this case, the spatial localization of biological causal factors) drive the methods, not vice versa. As multivariate methods began in true causation, so morphometrics began in the need to comprehend growth, ecophenotypy, speciation – the true biological causes and consequences of geometric form. As biometrics deterioriated over the decades into the manipulation of more and more naked matrices, its methods spoke less and less to the original purpose of shape description, namely, the explanation of shape change. Once the subject was restored to biological language, however ("What, exactly, are these forms *doing*?"), the methodological gap was filled in these few busy recent years.

The lesson for the broader context of applied statistical methodology, while the easier to state, surely is the more problematic in practice. *When an applied problem refuses, decade after decade, to submit to a broad assortment of assaults by analogy, consider carefully whether the problem has been misspecified.* Sixty years of competent biometrical explorations had failed to supply a statistical method that corresponded to Thompson's grids, about which every graduate student of mathematical biology yet dreamed. Time, then, to revert to fundamentals, to suggest that the grids were not the proper object of analysis at all: instead they are the irresistibly effective medium of display or communication of an analysis that had yet to be designed. Decade after decade, the dozen or so workers I have cited here circled around this ineluctable conclusion – *morphometrics must be the geometrically reified description of effects on geometric shape* – without ever realizing the crucial tool that was lacking: an algebraic formalization of the effects that the biologist wanted to understand. The breakthrough came with the multiple discovery of shape space for landmarks and the names for all the directions of that space. Thompson's error, like my own of 1978, was unusually subtle. It is

not that biological explanations report the evidence of the grids, but that the grids report the evidence of the explanations.

In stating the "moral" of my story so starkly, I intend to incite speculations in other branches of applied metrology that may be in difficulties at the present time. (Perhaps some are represented in the other chapters of this compendium.) In the morphometrics of outlines, for instance (to begin closest to home), there seems as yet to be no possibility of progress corresponding to what has happened in the last decade for landmark data. (For the *combination* of outline data with landmark locations, extending the finite-dimensional shape space constructed for landmarks, see Bookstein & Green, 1992a,b). It is time, perhaps, for a new statement of the problem: given the boundary of a biological object, perhaps the left ventricle of the human heart (or its silhouette, or a particular plane section), and given some information about homology across samples, what, exactly, do we want to be able to say? What kinds of explanations are we interested in? Variation of the phenomenon under study must be reduced to a finite-dimensional representation before multivariate statistics can apply. What finite list of descriptors do we have in mind, and how do we attach biological or clinical medical explanations to them? In medical image analysis, likewise, we are entrammeled in a riot of approaches to image processing and "reconstruction" without the driving force of a clear scientific question. What do we wish to learn, as biologists, from a medical image, or a heap of them, and what representation of the information content of the pixels will accede to the usual statistical tools, and what new tools are needed? In environmetric studies, in statistical ecology, and in many other fields where data is spatially or geometrically distributed, likewise the literature seems to me to be lacking in a methodology for tying reasonable scientific questions to geometrically distributed answers. We need a generalized language of "synoptic weather patterns" (themselves the invention of the same Francis Galton who discovered regression) for these more general models, and empirical hints about how to discover them in data in a manner allowing for a coherent statistical analysis and reporting. For an experiment along these lines, see Sampson *et al.*, 1991.

The triumph of modern multivariate statistical methods in fields arbitrarily far from their biometric origins has seriously distracted us from properly understanding the true meaning of these methods in the biological sciences. The meaning of statistical methods is inextricably bound up in what a community of scholars believe to be the meaning of their data (cf. Kuhn, 1959; Latour, 1987). The easy availability of matrix manipulations, and the ease with which they can lead to publications and tenure, is no substitute for an understanding of the nature of the tie between "the data" and the styles of explanation that actually drive the discipline in question. As the example of morphometrics indicates, there need be no mathematical model of a phenomenon (for instance, of skull growth), and yet the geometric dissection of the observed patterns of that phenomenon can

be effective and suggestive (Fig. 4A, B), as long as there is a satisfactory quantitative model of the *descriptive process itself*, the formalism of landmarks and deformations by which the patterns on the scientist's retina are converted into explanations.

The morphometrics of the synthesis supplies just such a model of *description*, a model that bridges Thompson's grids and the multivariate school's vectors by tying the landmark location data to the *report* of a difference. In the absence of one of these models or the other – a model for the phenomenon, or a model for the description of the phenomenon – I would not expect modern biostatistical methods to be of much use in other branches of quantitative biology. For instance, the well-known aversion of molecular biologists to statistical analysis – "if the data need statistics, the experiment was designed wrong"– is consistent with a digital logic of causes and effects orthogonal to the entire biometric tradition from Galton on; and, indeed, molecular biology has neither contributed anything to biometrics nor borrowed any techniques from us. By contrast, the morphometric synthesis of the 1980's drives biometrics straight back to its roots in the observation of organic forms. In my view, it is a major intellectual triumph within contemporary applied statistics, perhaps the most important of the last quarter-century: the perfect match of descriptive and inferential technique to a powerful classical mode of qualitative scientific intuition.

ACKNOWLEDGEMENT

Preparation of this paper was supported in part by NIH grants NS-26529 and GM-37251 to Fred L. Bookstein.

REFERENCES

BLACKITH, R., 1965. Morphometrics. Pp. 225-249. In T. H. Waterman and H. J. Morowitz, eds., *Theoretical and Mathematical Biology*. Blaisdell, New York.

BLACKITH, R. & R. REYMENT, 1971. *Multivariate Morphometrics*. Academic Press, New York.

BOOKSTEIN, F. L., 1978. The Measurement of Biological Shape and Shape Change. *Lecture Notes in Biomathematics*, v. 24. Springer-Verlag, New York.

BOOKSTEIN, F. L., 1982a. On the cephalometrics of skeletal change. *American Journal of Orthodontics*, 82: 177-198.

BOOKSTEIN, F. L., 1982b. Foundations of morphometrics. *Annual Reviews of Ecology and Systematics*, 13: 451-470.

BOOKSTEIN, F.L. 1984a. A statistical method for biological shape comparisons. *Journal of Theoretical Biology,* 107: 475-520.

BOOKSTEIN, F.L., 1984b. Tensor biometrics for changes in cranial shape. *Annals of Human Biology,* 11: 413-437 .

BOOKSTEIN, F.L., 1986. Size and shape spaces for landmark data in two dimensions. *Statistical Science,* 1: 181-242.

BOOKSTEIN, F.L., 1989a. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 11: 567-585.

BOOKSTEIN, F.L., 1989b. "Size and shape": a comment on semantics. *Systematic Zoology,* 38: 173-180.

BOOKSTEIN, F.L., 1991. *Morphometric Tools for Landmark Data.* Cambridge University Press, New York.

BOOKSTEIN, F. L., B. CHERNOFF, R. ELDER, J. HUMPHRIES, G. SMITH, & R. STRAUSS, 1985. *Morphometrics in Evolutionary Biology.* Academy of Natural Sciences of Philadelphia, Philadelphia.

BOOKSTEIN, F. L., & W. D. K. GREEN, 1992a. A feature space for edgels in images with landmarks. To appear in *SPIE Proceedings, vol. 1768.*

BOOKSTEIN, F. L., & W. D. K. GREEN, 1992b. Edge information at landmarks. To appear in *Proceedings of the Second International Conference on Visualization in Biomedical Computing, SPIE vol. 1808.*

BOYD, E., 1980. *Origins of the Study of Human Growth.* University of Oregon Health Sciences Center, Portland, OR.

BURNABY, T. P., 1966. Growth-invariant discriminant functions and generalized distances. *Biometrics,* 22: 96-110.

CORRUCCINI, R. S., 1981. Analytical techniques for Cartesian coordinate data with reference to the relationships between *Hylobates* and *Symphalangus. Systematic Zoology,* 30: 32-40.

DUNCAN, O. D., 1984. *Notes on Social Measurement: Historical and Critical.* Russell Sage Foundation, New York.

DÜRER, A., 1524. *Vier Bücher von Menschlicher Proportion.* Somewhere in Northern Europe.

GOODALL, C. R., 1983. *The statistical analysis of growth in two dimensions.* Doctoral dissertation, Department of Statistics, Harvard University.

GOODALL, C. R., 1991. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society* B53: 285-339.

GOODALL, C. R. & K. MARDIA, 1991. A geometric derivation of the shape density. *Advances in Applied Probability,* 23: 496-514.

HOPKINS, J. W., 1966. Some considerations in multivariate allometry. *Biometrics,* 22: 747-760.

HOTELLING, H., 1936. Relations between two sets of variables. *Biometrika,* 28: 321-377.

HUMPHRIES, J. M., F. BOOKSTEIN, B. CHERNOFF, G. SMITH, R. ELDER & S. POSS, 1981. Multivariate discrimination by shape in relation to size. *Systematic Zoology,* 30: 291-308.

HUXLEY, J., 1932. *Principles of Relative Growth.* Methuen, London.

JOLICOEUR, P., 1963. The multivariate generalization of the allometry equation. *Biometrics,* 19: 497-499.

KENDALL, D. G., 1984. Shape-manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society,* 16: 81-121.

KUHN, T. S., 1959. The function of measurement in modern physical science. Pp. 31-63 in H. Woolf, ed., *Quantification.* Bobbs-Merrill, Indianapolis.

LATOUR, B., 1987. *Science in Action.* Harvard University Press, Cambridge.

LEWIS, J. L., W. LEW & J. ZIMMERMAN, 1980. A nonhomogeneous anthropometric scaling method based on finite element principles. *Journal of Biomechanics,* 13: 815-824.

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

MACKENZIE, D. A., 1981. *Statistics in Britain, 1865-1930: The Social Construction of Scientific Knowledge.* Edinburgh University Press, Edinburgh.

MARDIA, K. V. & I. DRYDEN, 1989. The statistical analysis of shape data. *Biometrika,* 76: 271-282.

MOSIMANN, J. E., 1970. Size allometry: size and shape variables with characterizations of the log-normal and generalized gamma distributions. *Journal of the American Statistical Association,* 65: 930-945.

OXNARD, C. E., 1973. *Form and Pattern in Human Evolution.* University of Chicago Press, Chicago.

OXNARD, C. E., 1978. One biologist's view of morphometrics. *Annual Reviews of Ecology and Systematics,* 9: 219-241.

REYMENT, R. A., 1991. *Multivariate Palaeobiology.* Pergamon, Oxford.

RICHARDS, O. W. & A. C. KAVANAGH, 1943. The analysis of relative growth-gradients and changing form of growing organisms: illustrated by the tobacco leaf. *American Naturalist,* 77: 385-399.

ROHLF, F. J., 1986. The relationships among eigenshape analysis, Fourier analysis, and the analysis of coordinates. *Mathematical Geology,* 18: 845-854.

ROHLF, F. J., 1992. Relative warp analysis and an example of its application to mosquito wings. This volume.

ROHLF, F. J. & F. BOOKSTEIN, eds., 1990. *Proceedings of the Michigan Morphometrics Workshop.* University of Michigan Museums. Ann Arbor.

SAMPSON, P. D., F. BOOKSTEIN, S. LEWIS, C. HURLEY & P. GUTTORP, 1991. Computation and application of deformations for landmark data in morphometrics and environmetrics. Pp: 534-541. In E. Keramidas, ed., *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*. Interface Foundation of North America, Inc., Fairfax Station, VA.

SNEATH, P. H. A. & R. R. SOKAL, 1963. *Principles of Numerical Taxonomy*. W. H. Freeman, San Francisco.

SNEATH, P. H. A., 1967. Trend-surface analysis of transformation grids. *Journal of Zoology*, 151: 65-122.

THOMPSON, D'A. W., 1917. *On Growth and Form*. Macmillan, London.

WRIGHT, S., 1968. *Evolution and the Genetics of Populations. Vol. 1: Genetic and Biometric Foundations*. University of Chicago Press, Chicago.

ON THREE-DIMENSIONAL MORPHOMETRICS, AND ON THE IDENTIFICATION OF LANDMARK POINTS

V. LOUISE ROTH

Zoology Dept.
Duke University
BOX 90325
Durham, NC 27708-0325 USA
VLROTH@DUKEMVS.BITNET

CONTENTS

ABSTRACT

In morphometrics for comparative biology, the types of data collected and the manner in which measurements and landmarks are specified influence both the observations and the conclusions one can make. Hence, choices are best made deliberately, and with attention to hidden assumptions. In this paper I discuss (1) some of the consequences for morphometrics of working with two-dimensional representations of three-dimensional objects; (2) desiderata in the selection of landmark points and measurements; and (3) a way to evaluate the completeness of information captured in combinations of coordinate and linear distance data. This paper is a step in the direction of a primer on three-dimensional morphometrics, and illustrates ways in which evolutionary morphologists can deal more explicitly with assumptions about the spatial geometry of their biological subjects.

INTRODUCTION

We and other organisms live in a four-dimensional world: a world of three-dimensional morphological structures that participate in processes which add a fourth, or temporal dimension. At times it is convenient to conceptualize this world differently, with reference to more or to fewer dimensions: Hutchinson's (1958) n-dimensional hyperspace described ecological niches in a way that provides insight into community structure; classical multivariate morphometrics considers organismal morphology in terms of a morphospace of dimensionality defined by a set of n distance measurements; a single dimension –the height of a tree, the length of a tentacle– may represent most of the information relevant to a particular question of interest. In terms of our immediate experience, however, morphology is three-dimensional.

There are exceptions, in which most processes or variation are confined to two dimensions. Insect wings are doubled layers of cuticle whose aerodynamic properties can be considered to a first approximation to vary according to their shape and extent in two dimensions, and whose pigments distributed in the same plane serve as visual signals; leaves vary in shape largely but not exclusively within a single plane because of constraints related to their function as surfaces that collect light and exchange gases. But morphometric study in systematics and evolutionary morphology has not been limited to such objects.

Vertebrate skulls, snail shells, foraminiferan tests, and even fish bodies (bilaterally compressed and symmetrical as many of them may be) are quintessentially three-dimensional. Yet if one surveys the equipment available in most up-to-date laboratories for evolutionary morphometrics, the most prominent tools are video screens and digitizer pads, which are flat. The most recently developed techniques currently widely available to systematists for the analysis and comparison of biological shape are, with a few exceptions, based upon coordinate data of two dimensions (Bookstein *et al.*1985; Rohlf & Bookstein, 1990); "three-dimensional data" (but not "two", the apparent default) appears as a special entry in Bookstein's recent major opus (Bookstein, 1991). What many of us think of as sophisticated three-dimensional imaging techniques in clinical medicine –NMR, ultrasound, CT-scanning– generally present solid objects in series of two-dimensional slices. Techniques based upon "outline data" are a very powerful addition to the morphometric repertoire, especially for forms with relatively few landmark features (Lohmann & Schweitzer, 1990; Rohlf, 1990; Straney, 1990); yet 3-D objects do not in fact have outlines -they have surfaces and

contours. (As used here, "contour" refers to a closed curve- especially the outline of a 2D slice of a 3D object drawn onto the object's surface.) How is it that the two-dimensional representation of three-dimensional objects has become standard practice? Why does this seem natural, and why does it not strike us perpetually as a compromise?

One answer, I believe, lies in the fact that we are visual animals. Vision is a consequence of the neural integration of pattems of light projected onto the two-dimensional surfaces of retinas. The retina itself consists of several layers of cells, but the layers are involved in successive stages of processing of information impinging on a topologically two-dimensional array of receptors. We perceive an illusion of depth, the third dimension, in our visual images by tricks our brains play on us. Most humans (though not all individuals) have stereoscopic vision. From our two differently positioned eyes our brain receives two slightly different views of an object; it compares the two, and the disparity, or displacement, between these images is sensed by us as depth, conveying the impression that one object in the visual field is in front of or behind another. Depth can be perceived by a single eye or at distances too great for stereoscopy using other cues which are familiar to many artists: light reflected at edges of objects produces highlights that block from view the contours of objects positioned behind them; converging lines and atmospherically dulled colors recede into the distance, etc.

As natural as two-dimensional representations appear to us, our world is in fact 3-D. Comfortable as we may be with images projected onto a plane, such representations always selectively omit information. Unless we are consciously aware of this process of selection, we will be prone in our interpretations to error or bias.

In this paper I discuss (1) some of the consequences for morphometrics of working with two-dimensional representations of three-dimensional objects; (2) desiderata in the selection of landmark points and measurements; and (3) a way to evaluate the completeness of information captured in combinations of coordinate and linear distance data. This essay is presented at the time of the 500th anniversary of Spanish support for a now-famous attempt to demonstrate Earth's 3-dimensional geometry (in which shape, but not size, was correctly deduced), and of the approximately 108th anniversary of the publication of *Flatland: A Romance of Many Dimensions* (Abbott, 1884).

THREE DIMENSIONS INTO TWO

Orienting a specimen in space for image capture, and recording distance and coordinate data from the object for morphometric comparison, are related problems in three-dimensional geometry. For both problems it is important to know what constitutes the minimum essential information for unambiguously fixing the object in space. If one is given no prior information, to specify the position of any one landmark point in space requires three parameters, corresponding to one value for each of the three coordinate axes. Standard (x,y,z) coordinates are distances from

a defined set of axes that are linear and orthogonal, but in fact distances from any well-defined set of points (provided no three are collinear) or axes (provided none are coincident) can substitute. Typically, coordinate values have polarity -plus or minus directionality- so if similar conventions are applied to measurement data, they will be equally informative. With no additional prior information, to define the configuration of n points rigidly in space, one must take a minimum of 3 n measurements, distributed so that at least three terminate on each landmark point.

This rule of thumb is modified and manifest in various ways, according to the additional information provided in the contexts of different problems.

Orienting specimens

Imagine an object consisting of points, the locations of which are described in space by three coordinates whose frame of reference is defined with respect to a camera. By definition, the (x,y) plane runs parallel to the photoreceptive surface in the camera ("horizontal" and "vertical"), and the z axis ("depth") emerges perpendicular to that plane. We place the object in front of the camera in such a way that the camera receives a specified two-dimensional projection; i.e., our objective is for a particular plane defined within the object and by its own morphology to be in an (x,y) plane, parallel to the photoreceptive surface in the camera. Three points define a plane, so in principle any three non-collinear landmarks in the specimen can be used to orient it. **Three (z-coordinate) distances (one per landmark), measured with reference to the camera or a plane parallel to the picture plane, determine the projection of the specimen and thus determine the nature (but not the orientation within the plane) of the image itself.**

Once the image is captured, it can be shifted in position or reflected. This amounts to specifying an additional two coordinates for each of the original three landmarks. We are working with a rigid object, however, so fewer than this will suffice: Two coordinates for the first landmark pin the image in place at one point; an x-coordinate for a second landmark fixes the object in place with respect to the y-axis, but permits reflection across an axis parallel to x; a y-coordinate for the third landmark eliminates this remaining degree of freedom (provided neither of these last two coordinates is zero). The total information employed for this case of three landmarks is thus seven coordinate values or distances; by assuming our object is rigid we have in effect specified two more parameters (the respective distances between the first landmark and the other two), for a total of nine (=3n, as stated above).

Because organisms vary, and commonly lack perfect symmetry, it is often preferable to choose a reference plane (for example, a sagittal plane) representing the best fit through more than three points. Alternatively, various well-defined axes of symmetry can be used. For example, in orienting crania of squirrels for photography for morphometrics, I locate a point on the ventral midline just anterior

to the foramen magnum, and another just posterior to the incisors, which define a longitudinal axis I use in positioning the skull. A second set of axes runs perpendicular to this, extending between points corresponding in bilateral symmetry. (For greatest precision, it is best to select points of reference that are widely separated. Points far from a point of rotation are moved a greater distance with the same angular displacement than are more proximally situated points, so slight differences are amplified and more easily detected and controlled.) To photograph a dorsal view, I fix the anterior end of the longitudinal axis in place with plasticine clay, and the posterior end an equal (measured) distance above the working surface (which itself is level and parallel to the camera's picture plane). The skull is then still free to wobble in rotation about this longitudinal axis, so I use bilateral symmetry in the dorsal view to identify the position in which the skull will be fixed: I rotate the skull until the areas circumscribed by the two zygomatic arches appear the same through the camera. This is equivalent to specifying the distance (in the plane of projection) of a landmark on one zygomatic arch from the longitudinal midline, although it is achieved by balancing the positions of two points at opposite ends of an axis: a line (the bilateral axis) rotating in a plane about another (the longitudinal) axis has one degree of freedom, and its position can be specified with one measurement. Thus again, the plane of reference for the object is fixed rigidly in space by three parameters: in this example, by the two (z) distances for the longitudinal axis, and one for the zygomatic arches. (Note that the dorsal and ventral views produce identical planes of projection; a difference between them is only apparent when objects are not transparent, or as a consequence of parallax; see below.)

In choosing the points, axes, or a plane for orienting the specimen for image capture, one makes a decision that imposes a geometry upon all data that will be collected from that image. This decision on a reference plane is an especially important one if, as is currently common practice, subsequent analyses and comparisons work directly with two-dimensional coordinates, or with distance measurements obtained from these two-dimensional projections. An object can be reconstituted in three dimensions using a pair of images taken from different perspectives: In essence, 2D coordinates can be obtained for each landmark in one image, and the second image, by showing the relative displacement of each landmark from the new perspective, provides the third piece of information for a total of 3n parameters. In systematic practice, however, descriptions and comparisons currently tend to work within the context of a single planar representation at a time.

The tendency is natural, in working with two-dimensional representations, for one's conceptualization of forms and processes to become confined to that plane. Truss networks (Strauss & Bookstein, 1982), for example, were constructed with the understanding that relative positions of three points in a plane are rigidly specified by the lengths of the sides of the triangle they form. Biological landmarks, however, are free to shift in three, not two, dimensions. As finite-element methods recognize

(e.g., Cheverud & Richtsmeier, 1986; see also Bookstein, 1991), volumetric elements –tetrahedra, rather than triangles– may be the most appropriate elements to consider. In a planar representation, if the vertex of a triangle is observed to draw closer to the other two, we may in fact be witnessing either a shortening of the triangle (a change in the distances between vertices), or a foreshortening (a change in its orientation out of the parallel plane). Points of maximum curvature on an edge, the intersection of a biological feature such as a suture line in a shell or bone and the border of the form in profile, and entire "outlines" of globular or otherwise relatively featureless forms are all geometrically-defined features that can be useful in comparisons. Such features may not be readily identifiable on the three-dimensional object itself, however, for they depend not only on their geometric projection onto the plane, but also on how the plane of projection itself is defined.

When comparisons are made exclusively between two-dimensional projections, all geometrical descriptions and biological conclusions are made with reference to that plane. Choosing a plane of orientation constitutes an assertion of homology, and in defining all measurements with respect to it we make an assumption -or, at least, the language of our description suggests- that the plane itself does not vary or change. Sometimes the particular plane chosen is not very controversial: the midsagittal plane, identified with respect to any number of unpaired midline structures, is arguably homologous across most of the phyla of the Bilateria. A midpoint between the eyes, the anus, and root of the dorsal fin, for example, can be used to define this plane for any number of fishes. (Although in our real biological world of imperfect symmetry, slightly different planes will be defined by different choices of "midline" landmarks.) Even such highly conserved planes of symmetry can be inappropriate for some purposes, however. If this sagittal plane were to be used in a description of ontogeny in flounder, the data would suggest that larval metamorphosis involves a tremendous shift and rearrangement in the positions of fins and viscera. What in fact occurs can not be explicitly documented without reference to the third dimension. Disruption and change occurs in the larval symmetry itself, and the change is best described as a migration of the eyes themselves to a single side of the animal.

The effect of choosing a particular plane of reference in the case of the flounder is striking. But even in subtler examples, a two-dimensional projection represents a choice among a variety of potentially homologous planes or axes. Consequences vary. The longitudinal axis in the squirrel skull example described above could with equal justification be defined with respect to the axis of the basicranium, the tip of the nasal bones, the posterior lip of the foramen magnum, the occlusal plane, etc., producing different orientations, depending upon the shape of each skull (its basicranial flexure, the relative elongation of its rostrum, etc.). For any set of comparisons it will be essential to consider the possibility that a difference observed in the positions of landmarks explicitly measured on the image may in fact arise from change in the implicit frame of reference. **Important changes can occur in structures that are**

used in orienting specimens and determining their projections -structures that will not necessarily even be visible in the image itself.

When an image is captured, a commitment is made to a particular plane of projection. As a safeguard against the omission of important information, it is best to record three - rather than two-dimensional coordinates from landmarks, or to archive multiple views of the object so that 3D coordinates can be obtained (see, e.g., Grayson *et al.* 1988). Nevertheless, most of the techniques available to systematists, both for representation and for analysis, make use of only two dimensions at a time. Even if coordinate data ultimately must be projected into two dimensions for analysis or illustration, access to three-dimensional coordinates will allow one to vary the planes of projection, and examine the effect of each such choice.

Reconstructing 3 dimensions from 2-D images

The number of images necessary for capturing landmark positions in three dimensions depends upon the distribution of those landmark points over the object. To obtain three coordinate values from flat images, each point must be viewed from at least two angles, and each view must either be taken from specified orientations of the object, or show at least three additional points that are visible in other views (and can therefore be used to determine orientation). The problem then becomes one of photogrammetry (e.g. Slama, 1980). Our own visual systems obtain depth information by assessing (i) the convergence angle of the eyes (through proprioceptors in the ocular muscles) -this in essence establishes the relative positions of the two frames of reference; and (ii) the "retinal disparity", or disparity in the position of a single point on the two retinal images -the two images of a point coincide if the point is in the focal plane, and they diverge with changes in depth (La Prade *et al.* 1980). In photogrammetry, a height-to-base ratio (the ratio between the focal distance and the distance between the cameras) provides the same information as a convergence angle, and parallax displacement in the stereophotographs allows computation of the depth or elevation of a point out of the plane (La Prade *et al.* 1980). Special problems arise in connection with biological objects. When shapes are complex, and objects opaque, landmarks easily disappear from view. The set of views that suffice for one specimen may not be adequate for another one because of subtle variations in shape that cause landmarks to disappear behind bulges, other projecting features, or surface contours of the object. It may not be possible to determine what constitutes a sufficient number of viewpoints to use on a set of specimens until every specimen has been carefully examined.

Parallax may cause important distortions in the two-dimensional representation of landmark positions. An orthographic projection is an ideal in which, by definition, the mapping of all points follows strictly parallel lines. Objects at

effectively infinite focal distance fulfill this condition, but at close range, the angles of projection of widely separated points actually diverge widely. When an object is oriented obliquely in one, two, or three dimensions with respect to the camera plane, vanishing points (to which actually parallel lines converge) are introduced to the perspective (Williamson & Brill, 1990). A related problem arises with the use of X-ray images, whenever the size of the object is large relative to the distance from the X-ray source. For these reasons, further development both of (i) equipment (three-dimensional digitizers, the Reflex (TM) Microscope) that allows acquisition, and of (ii) analytical techniques (mapping and transformation) that allow direct *comparison* of three-dimensional coordinate data, will be especially welcome.

THE CHOICE OF FEATURES TO BE USED IN A COMPARISON

Comparison involves a sequence of selection processes. First, the objects to be compared are selected: the relevant taxa, and/or specimens, are identified, and the particular elements (skulls, shells, entire bodies) selected. Then particular aspects or features (size, as defined in a specific way, or shape, as represented or measured in some specific manner) are isolated and subjected to comparison. The method of comparison itself must be chosen.

No comparison of biological objects involves *complete* descriptions. The information extracted from a biological specimen and used in a comparison is by necessity a subset of everything that can be known or said about the object. As questions in biology arise within particular conceptual or disciplinary contexts, particular attributes or features (of the object, or of the object's form or size, etc.) are chosen for representation, codification, and comparison. Indeed perception itself involves the selection and selective organization and combination of unitary pieces of information, and as such it is analogous to comparison. In visual perception, the primary responses of photoreceptors to light or dark become translated, further along the chain of processing, into information about entire objects, their motion and contours. In morphometrics, sets of measurements or coordinate values are ultimately translated into information about shapes and differences between objects. Both perception and comparison involve an iterative series of processes of abstraction.

Choosing the features on a biological object to compare is an important step in the process of comparison. Bookstein (1990a: 219) has described landmarks as "the points at which one's explanations of biological processes are grounded," and has offered a classification and ranking of landmark types. Further generalization may be possible about what qualities are to be desired in a feature (be it a landmark position or other measurement) that is used in biological comparisons. I suggest that

1. Choice of a feature should be **repeatable**: A feature of morphology to which we refer in our quantitative descriptions should be well-defined; that is, it should be defined in such a way that it can be unambiguously re-located in the same place on the same specimen by another worker, or found uniquely on another specimen. Finding a feature on a new specimen -correctly identifying corresponding points despite differences in geometry- requires an appreciation for the full range of morphological variation that is likely to be encountered, and some sensitivity to the effects of a varying geometry on visibility in and projection into different reference planes.

2. If a set of features is to characterize an object, they should be well-distributed over it; that is, with respect to the anatomical regions and the questions of interest, they should be **comprehensive** (see Strauss & Bookstein, 1982; Bookstein *et al.* 1985). The data collected for each individual feature should also be comprehensive- e.g., they should specify important qualities of that feature unambiguously; hence the emphasis in this paper on full three-dimensionality.

3. The features should be **meaningful**: They should be relevant to a question of interest, and the methods of comparison should yield indices that usefully capture, clarify, or characterize relevant similarities and differences.

In phylogenetic analyses, the features we consider relevant are termed homologues- features for whose genetic and epigenetic basis we have some evidence of genealogical continuity (Roth, 1988, Van Valen, 1982). In purely biomechanical studies, homology may be irrelevant, and the biometrical information of greatest utility may be linear distances corresponding to lever arms, ratios of mechanical advantage, or dimensions that determine physical tolerances (e.g., minimum diameters, moments of area).

Repeatability (criterion #1) and relevance (#3) may be the bases for Bookstein's (1990a) preference for landmarks defined by local characteristics, such as the juxtaposition of different tissues, and they may account for his dissatisfaction with landmark points that are identified by their geometric relationships to other features (e.g., dimensional extremes, such as points of greatest width). With such features problems can arise because global features of geometry can be the combined effect of multiple locally-acting factors; two points defined geometrically in the same way on different specimens need not have any biological correspondence (and at times may not even fall on or within the specimen itself).

There are biometrical advantages to comparing the locations of landmark points on a form (Bookstein, 1990a). However, biological evidence suggests that at times one-to-one correspondence of single points fails:

> The morphologist, when comparing one organism with another, describes the differences between them point by point and "character" by "character" ....and he falls readily into the habit of thinking and talking of evolution as though it had proceeded on the lines of his own descriptions, point by point, and character by character. (Thompson, 1942: 1036)

Another level of description –of entire surface regions, or of volumetric elements, or of qualitative aspects of structures rather than structures themselves– may in some instances be most meaningful (Roth, 1984, 1991) and bring us closer to identifying the biological processes of interest. Hence the appeal and utility of methods of comparison that interpolate between landmark points, such as D'Arcy Thompson's transformation grids (Thompson, 1942; Bookstein, 1978), and the thin-plate spline (Bookstein, 1990b).

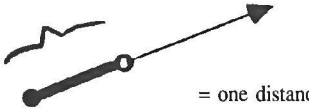
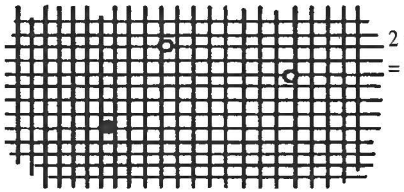
The recognition, and operational definition, of homologous points is a non- trivial problem (Jardine, 1969; Smith, 1990), and one not necessarily with unique solutions. As noted above, if one compares different objects, one inevitably encounters changing relationships between the component parts, and points defined in different ways may reflect different aspects of homology, and answer different questions. The appropriateness of any given set of landmark definitions is contingent –none is inherently preferable. For example, on any two bones, the scar on the humerus where the supinator longus inserts is a result of the same developmental process: the interaction of developing bone and muscle. Yet in terms of the processes that generate the overall gross morphology of a bone (such as major crests or constrictions), the positions of the scar on the two bones may not correspond at all. For biomechanical questions, a point of insertion may be of interest; but since a muscle may insert for quite a distance on the periosteum without penetrating the bone itself, the position of the scar may actually be misleading. Comparison of juvenile and adult bones presents particularly awkward problems since structures that are ossified in an adult may be cartilaginous in juveniles. A measurement of, for example, the longitudinal extent of osteogenesis (delimited by respectively the most proximal and distal points on a bone) allows the comparison of one aspect of bone development, but in terms of topography or of cell lineages within the bone, landmark points so defined are not homologous. Bone morphogenesis is a sufficiently complex epigenetic phenomenon, however, that information on cell lineages could be considered noise. Clearly, there is a plurality of different and valid approaches, and it is necessary not only to define but also to justify the choice of particular landmark points.

FUNDAMENTAL REQUIREMENTS FOR SPECIFYING THE LOCATIONS OF LANDMARK POINTS

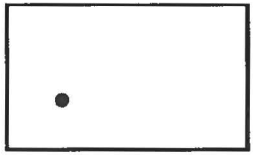
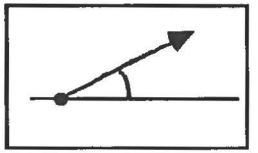
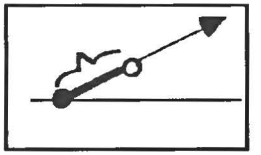
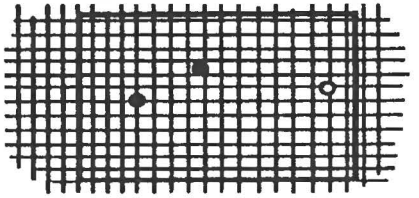
The comparison of three-dimensional shapes is a deceptively simple operation. A bulge in one portion of a bone may result from excess growth locally, from lessened growth in the surrounding regions, or from a shift in the relative positions of materials. A difference can usually be described multiple ways: "A is larger than B" sets B as the reference standard, and can be taken to imply that the condition of B is somehow primary, or perhaps primitive; "B is smaller than A" gives similar information, but with different implications; "A is large; B small"

Table 1

The minimum number of specified parameters required for n landmark points

Given	This is equivalent to being given	What is still needed
no prior information	zero information	the coordinates $(x_1, y_1, z_1), (x_2, y_2, z_2),... (x_n, y_n, z_n)$ **= 3n coordinates**
a fixed landmark point of reference	(a) the 3 coordinates for one point	the 3 coordinates for each of all (n-1) remaining points = (n - 1) points x 3 measurements **= 3n - 3 measurements**
a fixed point of reference, plus an axis of orientation	(a) plus (b) the 2 parameters required to specify the direction of a line in 3-D space  Total: 5 coordinates	the position of the second point along the axis of orientation  = one distance plus 3 coordinates for all remaining points = (n - 2) points x 3 measurements Total: 1 + 3(n - 2) = **3n - 5 measurements**
a fixed point of reference, plus a plane of reference	(a) plus (c) the 2 parameters required to specify the orientation of a plane Total: 5 coordinates	the positions of the second and third points in the plane  2 coordinates each = 4 measurements plus 3 coordinates for all remaining points = (n - 3) points x 3 measurements Total: 4 + 3 (n - 3) = **3n - 5 measurements**

[Table 1, cont'd]

Given	This is equivalent to being given	What is still needed
a fixed point of reference, plus an axis of orientation, plus a plane of reference	(a) plus (c)  plus (d) the one parameter required to specify the direction of a line within a plane  Total: 6 coordinates	the position of point #2 along the axis of orientation  plus the position of point #3 in the plane  = 2 coordinates plus 3 coordinates for all remaining points = (n - 3) points x 3 measurements Total: 3 + 3(n - 3) + **3n - 6 measurements**

Note: for each case, the numbers of measurements in the second and third columns sum to 3n.

The position of a point located *between* two others (i & j) and *collinear* with them can be uniquely specified by only two (rather than three) parameters: the distances of the point from each of the two points i and j. Therefore, for each set of three landmark reference points that are collinear, one fewer measurement than the standard number (given above in the third column) will be required.

is neutral in this regard, but suggests comparison with an absolute or external reference. A description of shape difference almost always (intentionally or not) involves implications of process.

Some of the most valuable contributions to the literature of comparative biology on size and shape historically preceded the formal technical and conceptual elaboration of morphometric techniques, and fall short of the ideal of providing an efficient and uniquely-determined specification of biological form. Even today, finances and logistics may limit the extent to which one can collect three-dimensional coordinates on landmarks for a particular comparison. If not optimal, it may be necessary or expedient to collect a combination of coordinate data and distance measurements. Both for planning studies today, and for making use of previous work, it is therefore useful to be able to evaluate a set of measurements for their ability to specify the positions of landmark points.

As stated earlier, minimally 3n parameters are required to fix the positions of n points in space, assuming they are arranged in such a way that at least three measurements (or coordinate values) terminate on each point. Ignoring any external frame of reference and concentrating upon the relative positions of landmark points within an object reduces the number of parameters to 3n-6; variable position and orientation of the object account for the six unspecified degrees of freedom. For a flat object, the number of parameters is 2n (or 2n-3 if position and orientation of the object relative to an observer in space are allowed to vary), so for a given number of reference points a three-dimensional object requires an additional n measurements for its specification.

In many instances, fewer than this number of measurements appears to suffice, because implicitly additional information about symmetry, axes of orientation, or identification of points is assumed (Table 1). For bilaterally symmetrical objects one may wish to specify the positional relationship between the two halves and then measure only one of the two sides. The additional information needed to describe the symmetrical half depends upon how many of the reference points on the measured half are shared between them. If three non-collinear points in the measured half of the object are in the sagittal plane, no additional data are necessary to generate the unmeasured half: since three fixed points define a plane about which no rotation is possible, the position and orientation are completely specified by the three points it shares with the half that has been measured, and its shape is presumed to be symmetrical to it. If only two points are shared, one additional measurement is necessary. The two points define an axis around which the two halves may rotate with respect to each other. If one additional distance–between any two symmetrical points from opposite sides –is taken, the positions of the two halves become fixed with respect to each other. With a single shared point, two additional measurements, between the two sides, should be taken, and with no reference points in common, one needs three measurements between the two sides.

It is possible, therefore, to erect the following general guidelines for taking measurements to specify the shape of an object: Once one has decided which n landmark points are of interest, one must decide upon (or decide to forego) a fixed point or a fixed orientation. One can then use Table 1 to determine the minimum total number of distance measurements or coordinate values required to specify their relative positions uniquely. The number is a minimum; where measurement error is a concern (e.g., usually; moreover, error propagates, if measurements are taken from other landmark points and not in relation to an external reference), some redundancy may be desirable (Rohlf & Archie, 1978). For n sufficiently large it is possible to define more measurements than are minimally needed to specify landmark positions, so a choice of measurements may be in order, but one needs to distribute the measurements in a way that provides all of the essential information. Except for points involved in fixing the frame of reference, and except for points that are collinear with two or more others, one needs at least three measurements terminating on each of the landmark points (and, as stated earlier, three distances, to points whose positions are already specified, provide the same amount of information as the specification of positions along the coordinate axes). For each set of three points that are collinear, one fewer measurement than the standard number is required, since a point located intermediate between two others and collinear with them requires only the distances from each to fix its position. By way of illustration, in Table 2 (see accompanying Fig. 1) I evaluate a set of measurements I took on tibias for morphometric comparisons of elephants (Roth, 1982; 1992).

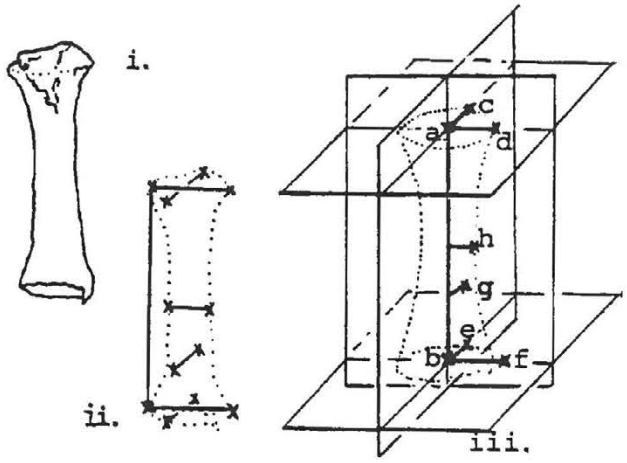


Fig. 1

i. Right tibia of an elephantid, anterior view. ii. Endpoints of measurements taken on tibias, shown diagramatically. iii. Planes of symmetry implicitly suggested by this set of measurements. See Table 2 for further discussion.

The precise specification of shape may not be the only objective of measurement: the overall proportions or dimensions of an object may be of more interest than the exact positions of particular reference points. When suture lines are wavy, for example, or boney projections are irregular, it is not difficult for the precision of one's measurements to exceed the precision with which one can define a reference point. One may be satisfied with the approximation of shape that a relatively small number of measurements provides, or to make comparisons one may be constrained to use a particular set that another author has used

Table 2

Evaluation of a set of measurements taken on tibias (Roth, 1982; 1992). See accompanying Figure 1.

The following measurements were initially judged informative and expedient:

1. Length of diaphysis
2. Maximum anterior-posterior diameter of the proximal end
3. Maximum transverse diameter of the proximal end
4. Minimum anterior-posterior diameter of diaphysis
5. Minimum transverse diameter of diaphysis
6. Maximum anterior-posterior diameter of distal end
7. Maximum transverse diameter of distal end.

In an optimal world, with the additional experience I have now accumulated from years of staring at elephant tibias, applying the criteria of repeatability, comprehensiveness, and relevance described in this section, I might now define a somewhat different set of landmarks and comparisons to characterize variation in tibial morphology. Elephant specimens are few, however, and widely dispersed among museums (and on the basis of size alone, n=1 is a large sample of elephants), so some opportunism is necessary, and it is desirable to make the most of available data. This set of measurements is reasonably representative of those traditionally taken for morphometric study of vertebrate long bones.

In taking the measurements I used twelve reference points (see Fig. 1, ii). To specify the locations of these points rigidly one would need a fixed point and a specified orientation for the bone, plus (see Table 1) 3n-6, or 30 measurements. If, however, certain simplifying assumptions (idealizations of the shape) are made, this set of 7 measurements comes close to describing the locations of the reference points completely.

In taking the seven measurements to be adequate or nearly so, one is implicitly assuming that it is possible to (A) define two planes of symmetry, the transverse and the anteroposterior, which are perpendicular to each other and intersect along the central axis of the bone; (B) consider all points defined to be at the proximal end to lie in a single plane perpendicular to the central axis; and (C) consider all points at the distal end similarly. With assumption (A) we effectively reduce the number of reference points to 8 (by symmetry, description of one quadrant of the bone is sufficient; see Fig. 1, iii). If all distal points lie in a single plane (ditto for proximal points), and if the antero-posterior and transverse planes are predetermined, the localization of points c,d,e, and f requires only a single measurement apiece. Points g and h by definition lie in the anteroposterior and transverse planes, respectively, so only two coordinates are needed for the location of each; and one distance (a-b) determines the relative locations of points a and b (and consequently the locations of proximal and distal planes). Therefore, if the stated assumptions are not too great a violation of reality, 4 + (2 x 2) + 1 = 9 measurements should suffice: to complete the set, one need only add the distances from the end of the bone at which measurements #4 and #5 (above) are taken, to position points g and h along the vertical axis.

previously. Whatever features, landmarks, or comparisons one ultimately chooses, the best outcome can be expected when choices are made deliberately, with attention to hidden assumptions, and when they are a product of evaluation, not prescription.

ACKNOWLEDGEMENTS

This paper is derived from a chapter of a dissertation presented a decade ago in partial fulfillment of a PhD degree at Yale University. Its updating and more formal presentation were stimulated by (U.S.) NSF-sponsored second Morphometrics Workshop at SUNY (Stony Brook), 1990, with additional financial support provided by NSF grant DEB 85-16818. I thank Hugh Crenshaw and John Mercer for helpful discussion, and Les Marcus, Fred Bookstein and Jim Rohlf for comments on the manuscript. John reminded me of the important anniversaries.

REFERENCES

ABBOTT, E. A., 1884. *Flatland: A Romance of Many Dimensions.* [1963 reprinting by Barnes & Noble, New York; 108 pp.]

BOOKSTEIN, F.L., B. CHERNOFF, R. ELDER, J. HUMPHRIES, G. SMITH & R. STRAUSS, 1985. *Morphometrics in Evolutionary Biology.* The Academy of Natural Science, Philadelphia, Spec. Publ. No. 15, 277 pp.

BOOKSTEIN, F.L., 1978. The Measurement of Biological Shape and Shape Change. *Lecture Notes in Biomathematics*, vol. 24. Berlin: Springer.

– 1990a. Introduction to methods for landmark data. Pp. 215-235. In *Proceedings of the Michigan Morphometrics Workshop* (F.J. Rohlf & F.L. Bookstein, eds.), Special Publication No. 2, University of Michigan Museum of Zoology.

– 1990b. Higher order features of shape. Pp. 237-250. In *Proceedings of the Michigan Morphometrics Workshop* (F.J. Rohlf & F.L. Bookstein, eds.), Special Publication No. 2, University of Michigan Museum of Zoology.

– 1991. *Morphometric Tools for Landmark Data.* Cambridge U. Press.

CHEVERUD, J.M. & J. RICHTSMEIER, 1986. Finite-element scaling applied to sexual dimorphism in rhesus macaque *(Macaca mulatta)* facial growth. *Systematic Zoology,* 35: 381-399.

GRAYSON, B., C. CUTTING, F.L. BOOKSTEIN, H.-C. KIM & J. MCCARTHY, 1988. The three-dimensional cephalogram: theory, technique, and clinical application. *American Journal of Orthodontics,* 94: 327-337.

HUTCHINSON, G.E., 1958. *Concluding remarks.* Cold Spring Harbor *Symposium Quantitative Biology,* 22: 415-427.

JARDINE, N., 1969. The observational and theoretical components of homology: a study based on the morphology of the dermal skull-roofs of rhipidistean fishes. *Biological Journal of the Linnean Society,* 1: 327-361.

LA PRADE, G.L., S.J. BRIGGS, R.J. FARRELL & E.S. LEONARDO, 1980. Stereoscopy. Chapter X (pp. 519-544) in *Manual of Photogrammetry*, 4th ed. (C.C. Slama, C. Theurer, & S. W. Henriksen, eds.) American Society of Photogrammetry, Falls Church VA.

LOHMANN, G.P. & P.N. SCHWEITZER, 1990. On eigenshape analysis. Pp.147-166. In *Proceedings of the Michigan Morphometrics Workshop* (F.J. Rohlf & F.L. Bookstein, eds.), Special Publication No. 2, University of Michigan Museum of Zoology.

ROHLF, F.J., 1990. Fitting curves to outlines. Pp.167-178. In *Proceedings of the Michigan Morphometrics Workshop* (F.J. Rohlf & F.L. Bookstein, eds.), Special Publication No. 2, University of Michigan Museum of Zoology. Rohlf,

ROHLF, F.J., & J.W. ARCHIE, 1978. Least-squares mapping using interpoint distances. *Ecology,* 59: 126-132.

ROHLF, F.J. & F.L. BOOKSTEIN (eds.), 1990. *Proceedings of the Michigan Morphometrics Workshop*, Special Publication No. 2, University of Michigan Museum of Zoology.

ROTH, V.L., 1982. *Dwarf mammoths of the Santa Barbara, California Channel Islands: Size, shape, development, and evolution.* Ph.D. dissertation, Yale University, 277 pp. Michigan Microfilms, Ann Arbor.

– 1984. On homology. *Biological Journal of the Linnean Society,* 22: 13-29.

– 1988. The biological basis of homology. Pp. 1-26. In *Ontogeny and Systematics* (C.J. Humphries, ed.), Columbia U. Press.

– 1991. Homology and hierarchies: problems solved and unresolved. *Journal Evolutionary Biology,* 4: 167-194.

– (1992). Dwarfism and variability in the Santa Rosa Island mammoth: An interspecific comparison of limb-bone sizes and shapes in elephants. In: *Recent Advances in Channel Islands Research.* Proceedings of the Third California Islands Symposium (F.G. Hochberg, ed.) Santa Barbara Museum of Natural History.

SLAMA, C.C. (ed.), 1980. *Manual of Photogrammetry*, 4th ed. American Society of Photogrammetry, Falls Church VA.

SMITH, G.R., 1990. Homology in morphometrics and phylogenetics. Pp. 325-338. In *Proceedings of the Michigan Morphometrics Workshop* (F.J. Rohlf & F.L. Bookstein, eds.), Special Publication No. 2, University of Michigan Museum of Zoology.

STRANEY, D.L. 1990. Median axis methods in morphometrics. Pp. 179-200. In *Proceedings of the Michigan Morphometrics Workshop* (F.J. Rohlf & F.L. Bookstein, eds.), Special Publication No. 2, University of Michigan Museum of Zoology.

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

THOMPSON, D.W., 1942. *On Growth and Form*. Cambridge U. Press 1116 pp.

VAN VALEN, L., 1982. Homology and causes. *Journal of Morphology* 173: 305-312.

WILLIAMSON, J.R. & M. H. BRILL, 1990. *Dimensional Analysis through Perspective*. Kendall Hunt, Dubuque.

PART TWO
DATA ACQUISITION

BUILDING YOUR OWN MACHINE IMAGE SYSTEM FOR MORPHOMETRIC ANALYSIS: A USER POINT OF VIEW

JOSE M. BECERRA, ELISA BELLO & ANTONIO GARCIA-VALDECASAS[1]

Museo Nacional Ciencias Naturales
José Gutiérrez Abascal, 2
Madrid - 28006 (Spain)
MCNJM14@CC.CSIC.ES; MCNEB18@CC.CSIC.ES;
MCNHV00@CC.CSIC.ES

[1] Alphabetical authorship order.

CONTENTS

ABSTRACT

Some of the hardware and software available for video image acquisition for 2D morphometric analysis on the PC is reviewed, with a special emphasis on cost/effect relation. We consider hardware for image acquisition from video cameras, still video, and digital cameras to secondary machinery like CDs and VCRs. Images that are in video format (analog) need to be digitized for PC processing. This is attainable through frame grabbers. We report experience with video frame grabbers from Imaging Technologies and on the less expensive and newer video-VGA cards. Resolution evaluations in different combinations of cameras/monitors and grabbers are presented. Morphosys, MTV and Java data acquisition software are compared in terms of accuracy. A short comment on storage, compression and format translation is included.

INTRODUCTION

There is much recent literature on automatic capture and treatment of images for use in systematic biology, especially for the repetitive acquisition of morphometric data (see Fink, 1987, 1990; Macleod, 1990; Rohlf, 1990a; Meacham, 1992). A very useful "practical primer" on image digitizing is Lindley, 1991, where the basic notions can be found. Jähne (1991) is a more technical up to date treatment. Meacham's paper complements our objectives, as it deals with critical aspects of object illumination and optical deformation of the image. Nevertheless, when we decided to procure an automatic image system for two dimensional (2D) morphometric analysis, we had to make practical decisions, whose answers were not easy to find in the literature. We think that a summary of our experience will prove useful to other systematists.

We decided not to buy an "off the shelf" assembled system, and instead built our own. Either choice requires much thought. There are many systems on the market, and they should be evaluated not in terms of what they offer for the price, but rather in terms of what you will actually use in relation to the price. Many of the already assembled systems offer multiple functions and great flexibility, but most of this is of little or no use in systematic work. A wide range of equipments is described in Data Sources (Anonymous, 1992).

If you are going to build your own system, there is no easy solution. You will vacillate between numbers of options and economy, and there is an inverse relation between the two. The more options and flexibility you want, the more you must be prepared to spend. Some sort of balance is possible, given your budget, and we discuss the alternative systems we have considered and tested.

MACHINE IMAGE SYSTEMS: AN OVERVIEW

Fig. I displays a general conception of the steps in the capture and analysis of images in the morphometrics area. Rohlf's (1990b) distinction between data acquisition, feature extraction and morphometric analysis has been simplified somewhat as his first two steps are here labelled under Data Acquisition. Only this step will be dealt with here.

The flow of the digitatization process is as follows:

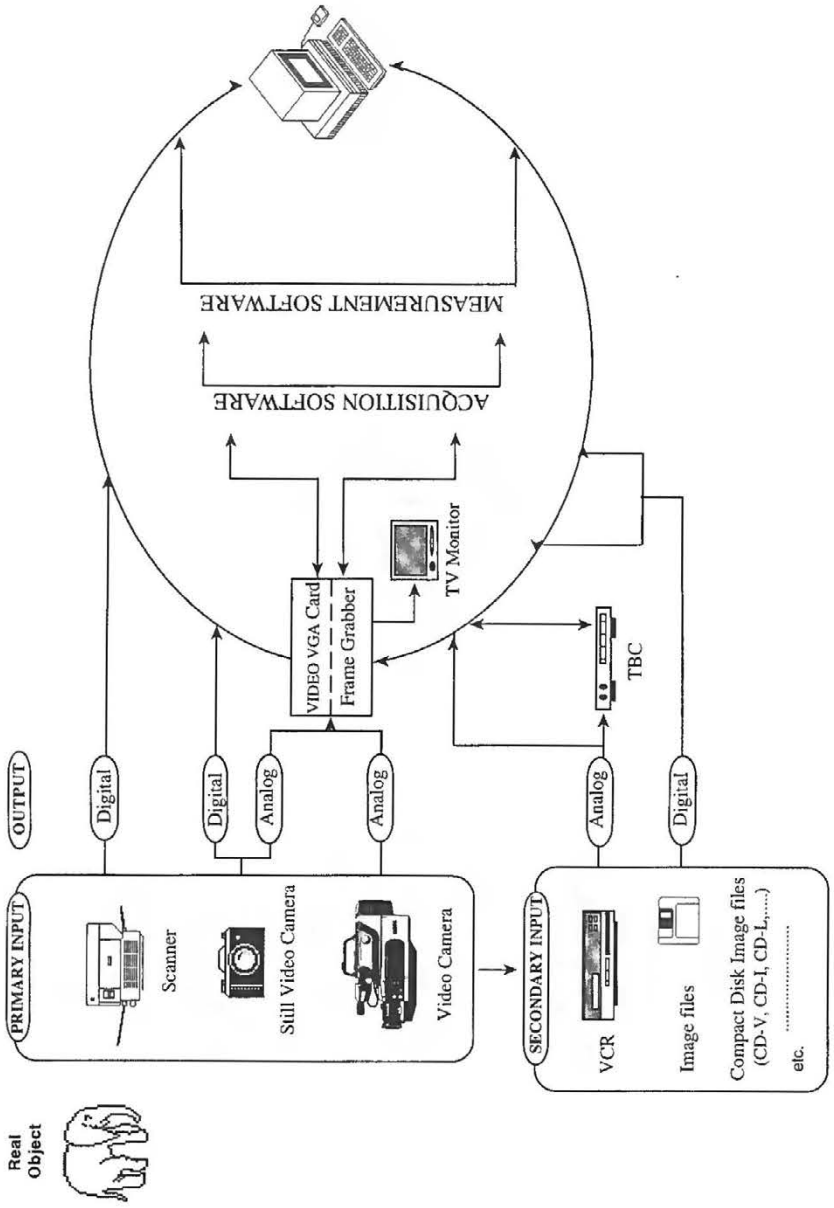


Fig. 1

General flow of the information in an image acquisition system and the elements that build up that system.

The real object is acquired by a sensor. In some cases (analog still video cameras and video cameras) the sensor is read out and an analog video signal is created. In some others, the system sensor generates a digital signal that can be used directly by the computer.

A video signal is usually displayed as frames that are built of video lines, 625 in the PAL and SECAM standards and 525 in the NTSC standard. Video lines are assembled into fields and two fields make a frame. One has the odd video lines and the other the even video lines. The frame is the interlacing of the two fields. To make the analog signal available to the computer, we need an analog-to-digital converter card (either a grabber or a Video to VGA board). The video signal is sampled and then translated into digital format and stored in computer files (Luther, 1991).

CRITERIA FOR A MACHINE IMAGE SYSTEM

All the parts one should have in order to operate an automatic image system can be described in two parts: 1) hardware; and 2) software. These are not completely independent. As this paper is restricted to personal computers, all considerations depend on a choice between the IBM PC standard and its clones, and the MacIntosh. No further mention will be made of dedicated workstations for image treatment, such as UNIX based systems and others, for which the Data Sources reference mentioned above could be useful. Most programs for the morphometric analysis described in the series of workshop proceedings, to which this volume belongs, are available only on IBM PCs and clones. Therefore, all of our discussion will consider only the IBM PC environment (hereafter referred to as just PC). However, a mixed system - using a Mac for data acquisition and a PC for data analysis is possible.

What were the criteria for our choice of an image system? Budget was a constraint, but for the same amount of money a knowledgeable purchaser can get more features than an inexperienced one. Commercial sales people can be helpful, but for one or another reason, they often offer you things that you don't really need or that can become a luxury if funds are scarce. When acquiring an image system for morphometric analysis, you should already know the kind of data processing you would like to do. Then obtaining the right software is easily discernible. Lots of extra functions and routines can only add to the price of your system. If in the future you need a new software tool, then that will be the appropriate time to buy it. Software and hardware costs tend to lower, and paying today for something you will only use tomorrow is bad practice. That is the main reason why the only software reviewed in this paper are MorphoSys, MeasurementTV (MTV) and Java (we have unsuccessfully tried to obtain a copy of CODA). All three have been developed with the morphometrician in mind.

Other software can do morphometric acquisition and much more, but you will probably not use the other features unless you have very specific needs.

What about selecting and testing hardware? This is most confusing for beginners. There are a variety of cameras, digitizers and other tools, not to mention display and TV monitors and models of computer. In order to resolve this problem we had to compromise. We decided not to test each component separately. Tests for some components require very sophisticated and expensive apparatus, and they are not worth the investment for a single installation. What we were really interested in was total system performance. So, apart from considering single items when appropriate, we tested whole systems. Before covering equipment and components we will discuss some general concepts.

Two main things are relevant for evaluating the adequacy of a machine image system: resolution and accuracy. A general definition of resolution is the capacity to discriminate between two near points. In terms of a TV system, horizontal resolution is defined as the number of black and white vertical lines that can be reproduced at a distance corresponding to the raster height (Luther, 1991). 450 lines of horizontal resolution correspond to 225 black and 225 white vertical lines. Vertical resolution depends on the number of horizontal TV lines (525 or 625 but not all those lines are video active lines. See Mcleod, 1990, for a more detailed comment). That is the main reason why European standard TV systems (PAL, SECAM), all else being equal, have better vertical resolution than American and Japanese (NTSC). Of course, all this is going to change in the near future with the emerging new high definition TV. Resolution in frame grabbers is also dictated by the sampling frequency in digitation, and this is dependent on the internal clock of the frame grabber. As the duration of the TV signal is fixed for every TV standard, sampling frequency determines the number of points per video line taken from the analog signal. For example, PAL TV systems (there are several PALs) have 625 lines per frame and 25 frames per second. Each line has a duration of 64 μseg of which, when synchronism information is deleted, gives approximately 52 μseg of useful information for digitization. If these 52 μseg are sampled with a 10 MHz clock (and this means 10 million samples every second) we get 512 pixels per line. If the sampling is done with a 12.5 MHz then 640 pixels are obtained.

Accuracy (Sokal & Rohlf, 1981) is the closeness of a measured or computed value to its true value. It is related to precision, which is the closeness of repeated measurements to the same value. The situation for image systems is rather frustrating, because although all measurement and engineering devices include information on measurement error, we have yet to see any estimation of this error in any of the completely assembled systems or software that we have looked at.

One method of finding horizontal and vertical resolution, is to use a standard TV chart (Fig. 2). Then it is easy using one to see how many horizontal and

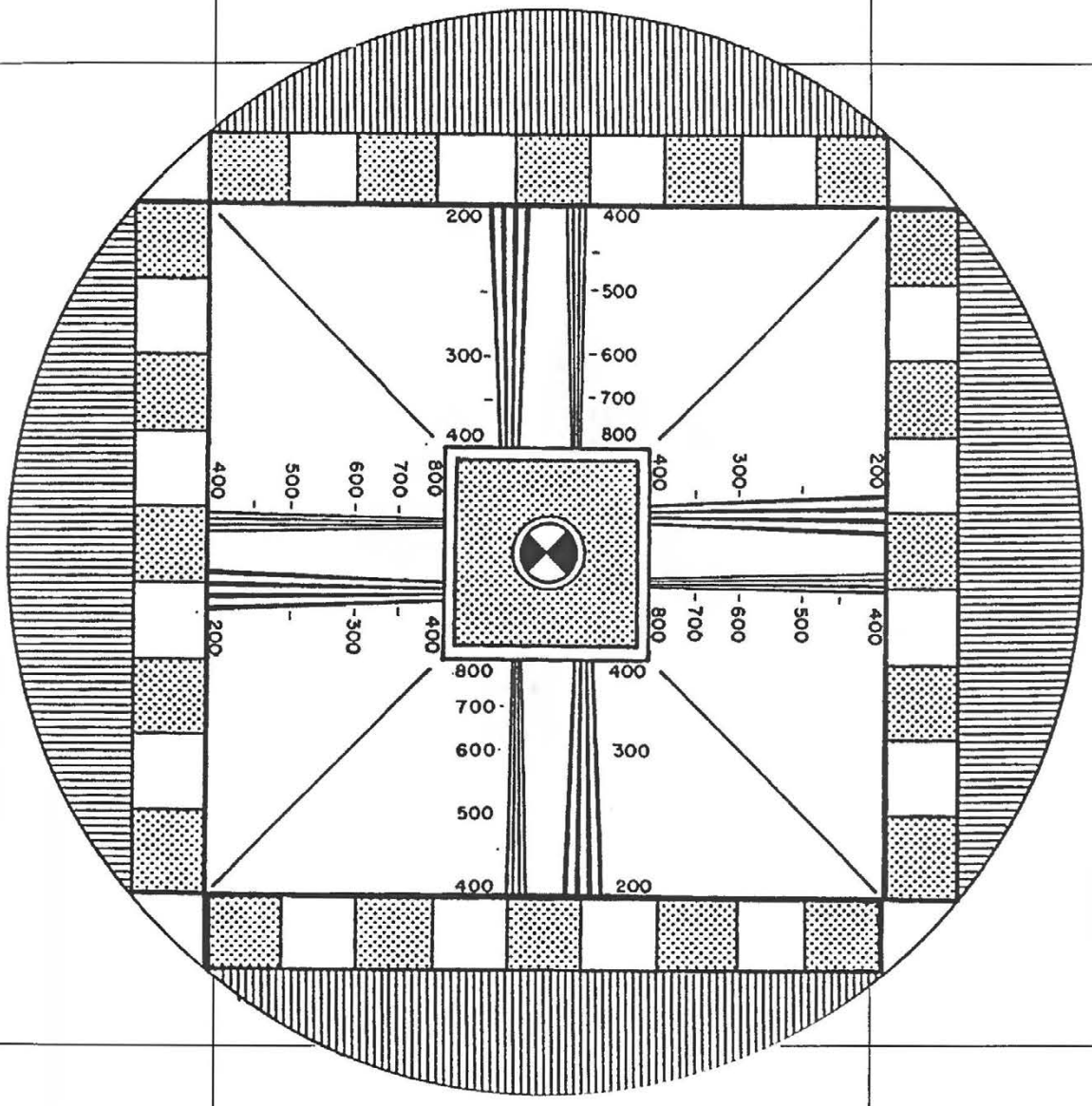


Fig. 2

Standard TV chart (It may be used for approximate evaluation of resolution).

vertical lines the system is able to discriminate. We describe the results of such a test (see below) for our available equipment, after covering the components of the system.

As other practical criteria, it can be of interest to consider some of the following possibilities: same system able to work with stereo or light microscopes, and by itself with different kinds of objectives; able to be taken to the field or other museums for laboratory work and recording.

HARDWARE

The minimal configuration you need for a machine image system is:

a) Image input devices. They can be divided into primary sources devices, that is, those that take images from the real world (whatever that means!). They could be a scanner, a video camera, a still video or still digital camera (we exclude from our discussion digital calipers and digitizing tables, as well as TV devices which can not digitize images). A scanner could benefit from a preprocessing, like scanning images from a previously photographed or drawn object. Secondary sources are those that temporarily store the images from a primary source. They can be an image saved as a file, in VCRs, laser disks, and other devices. The main point here is that when you go from primary sources to secondary sources you always lose resolution. Hybrid systems are always possible, for instance, a camera or a VCR and a TV pointing device, that takes data -mainly coordinate data points- from the TV image and creates a data file in the PC. Measurements may be made from images previously stored on a VCR or magnetic disk.

b) Usually, but not always, an additional device, a frame grabber or digitizer. Frame grabbers are necessary if your image input system produces an analog video signal. These analog signals must be converted to digital form before processing by the computer. Video and still video cameras belong to this type of equipment. Scanners and digital cameras give a digitized image as an already built-in function. Frame grabbers usually digitize video in "real time", that is 25 or 30 frames per second. Lower priced video-VGA boards digitize and display on a computer EGA or VGA monitor.

c) A reasonably fast PC (386 or 486 CPU) with VGA and large hard drive to store images. A math-coprocessor is desirable, and if Microsoft Windows is to be used, at least 4 MBytes of RAM.

d) A TV monitor is common for most digitizing boards, but as was pointed out above, some digitizing boards use the computer monitor. Multiscan monitors can serve as both a VGA and video monitor, and are switchable.

There are other additional devices that could be useful, but they will be mentioned in context. We will discuss each of these components in more detail summarizing our experiences with some of them.

Video cameras

Video cameras are the most common device for image input. As mentioned above they use NTSC, PAL, or SECAM standards. It is important that other parts of the equipment be compatible with the standard your camera is using. They also may be black and white, or color. The camera may also have the capability of storing images on video-tape and in this case is called a cam-corder. If the camera does not have this capability, it can be called an "in situ" camera. The cam-corder may be classified as amateur or professional, depending on features, manufacturing care, and price. Amateur video camera recorders have up to 450 TV lines of horizontal "resolution" (and the number is going up). More expensive professional video camera recorders go beyond 700 TV lines. In-situ cameras can be attached to different kinds of lenses, or to microscopes and stereo microscopes. With a C mount or other compatible mounting ring, they can be used with camera lenses. Among amateurs cameras only the Canon EX1 camcorder (L1 in the US) has detachable lenses, is compatible with the Canon EOS series of camera lenses, and can be adapted to a microscope with a custom adapter (about $60-70).

Light sensitivity of the camera is an important feature, as it can greatly facilitate the crucial step of taking the image. Recent amateur cameras come with very good sensitivity, from one lux onwards. Light sensitivity is a different question from proper illumination of the object, a crucial element when working with digitized images. Problems related with object illumination are dealt with in Meacham (1992).

One reason for considering built in recording capabilities is that many systematists frequently travel to other institutions to record data on specimens housed there. It seems sensible to "tape" the information with the same camera you use in your home institution, in order to avoid any transfer among tapes or video systems that can only degrade the resolution and quality of the image. There is the added value of obtaining images of live specimens in the field or laboratory. Resolution of camcorders is in the following order: VHS = 8 mm < SVHS which is the same as Hi 8. Another useful feature when using camcorders is remote control, which allows remote focusing and frame by frame viewing, among other features.

We have not had any experience with any professional video recording formats like Umatic, Betacam, etc.

Another matter, is the format of the video signal that the video camera sends. Composite seems to be a cheap way of coding the video signal, but usually does not lead to very good resolution. S-Video gives better results. These are the two most common ways of encoding color TV signals in amateur video cameras. Other formats that can be found for this market in the near future are Component, RGB and YIQ. It is important to know which of these formats your camera has, as the frame grabber has to be able to input the video signal in one (or several) of these formats. In our experience, we have found that those cameras that have composite and another optional video signal output like RGB or S-Video, have poorer resolution in the composite than in the other one. We have looked at the Sony DXC-930 (a 3-CCD in-situ camera) that has 720 lines of horizontal "resolution" (l.h.r.). This high number of lines is only obtainable in the RGB output. The composite output of this camera gives values below 500 l.h.r. Similar results are found with top end camcorders like the Sony V5000E or the Canon EX1. See below for resolution comparisons. It seems that sending high resolution composite signals is more expensive than doing the same in the other formats. Video cameras could include several video output formats, but the advertised resolution may be found in only one of the video outputs.

A final comment on optics. It is a critical part of any video camera. Lower priced camcorders do not have good optics, and in some cases the lenses are plastic. Higher quality cameras can have removable optics that usually can be sustituted by an ad hoc adapter to other inexpensive optics. In this way we have been able to attach the Canon EX1 to different types of microscopes with a special hand-made attachment (see above).

A choice between an "in situ" camera and camcorder then depends on the following differences in features. Color on most camcorders vs black and white on in situ cameras, built in optics on camcorders (except the EX1) vs wide flexibility for the in situ camera, in situ camera must be used with a digitizer (or at least a VCR) while a camcorder has its own tape recorder and player.

A few hints to select a video camera follow:

1) Decide the resolution you would like in terms of the compatibility with the rest of your system.
2) Look at in situ cameras in the range of your budget, and good quality camcorders. Make a tabular comparison including the following items: Resolution (in terms of number of horizontal lines), formats of video signal, taping quality and availability of the optics used. Some additional items are availability and cost of remote control.
3) Check that the video signal you are going to use with your frame grabber has the resolution you decided on (1), and is in the format that your frame grabber will accept.

4) Decide if you are you going to use the same camera for different purposes - that is, by itself, or in conjunction with a microscope for example.
5) Once you have decided which cameras fit your budget and your quality demands, if at all possible test them with the resolution chart.

We think that the available amateur camcorders give adequate resolution in relation to digitizing boards and are reasonably priced. Get a black and white in situ camera, unless color is absolutely necessary and you can afford it.

Cameras tested in this study has been: the Sony F-555E and V5000E, the Canon EX1 (all three color camcorders), and the Cohu 4815/2000 (B&W). Some other cameras have been looked at in a more casual way.

Still video and digital cameras

We briefly discuss still video cameras, whose potential for use in morphometrics is considerable, but models with adequate resolution are still prohibitively expensive when compared to the best of amateur camcorders and affordable "in situ" cameras. They derive their name from the fact that they take one image at a time, and then save it either in analog or digital format.

Digital still cameras typically store the images in a chip inside the camera (solid-state memory), and the number of pictures you can take are limited by the size of memory of that chip, as the amount of memory taken per image is fixed (and kept low through compression techniques). For instance, DYCAM Model 1 camera (whose technology has been bought by LOGITECH and now manufactured under the name of FOTOMAN) can store 32 pictures at 376x240 pixel and 256-gray scale. Once you have exhausted your camera memory capacity, adequate software (included in the camera set) must be used to tranfer (via a serial link) the images to your computer or laptop before beginning again (unless you want to rewrite a previous stored picture!). Unfortunately, the poor quality of the optics of this camera makes it inadequate for morphometric work. Canon is working on a digital still camera with 1,300,000 pixels with the EOS optics for the amateur market, and it is possible that it will be available soon, despite its recent upgrade of the Ion (see below). Top digital camera models like the Rollei Digital Scanback with 32 million pixels in 6x6 cm format or the Kodak Digital Camera System -DCS 200- with 1,54 million pixels in 24x36 mm format based on a Nikon N8008a or the Megaplus XRC with 2,680 x 1,035 pixels are prohibitively expensive.

Analog video cameras transfer the image to a floppy disk in the camera as an analog signal. Additional equipment - a digitizer - is necessary to make your picture available to your computer. Typically, you can store 50 pictures (typically only 25 frames and 50 fields) in a 2" floppy disk and you can use as many disks as you want. The Canon Ion RC 260 still video camera (Xapshot in the USA),

records one field of 320 TV lines, but its lens has serious limitations. It is a fixed focus lens able to focus from 1 m to infinity and from 30 cm in macro position. The camera is not designed to be adapted to microscopes or stereomicroscopes. The Canon Ion RC 560 is improved over the 260 in that it has 450 horizontal TV lines and a much better optical system, similar to rangefinder 35 mm cameras,that includes autofocus and a 3x zoom. Battery life is a possible drawback.

The Sony Mavica, that can be found on the US market, is not available in Europe, and it seems it will not be (Sony Spain dixit). Nevertheless, it has been advertised as a high resolution still video camera: 500 horizontal lines and 3 CCD. It has a typical reflex optical system.

Scanners

Scanners are optical devices that usually work with photographs and drawings, i.e. images on flat surfaces. They produce an image file in one or more of the several formats for images files (e.g. GIF, PCX, TIFF, etc). We have not been able to test a reasonable range of scanners so the reader is referred to the literature. See Beale and Cavuoto, 1991, or Glover, 1990 for a general overview.

Digitizers

Video digitizers digitize analog video signal and then using adequate software it can be posible to store and display the digitized image. A first subdivision is made between those that digitize the video signal on the board (grabbers) and those that are complementary to the VGA board of your computer (Video to VGA boards).

Among the grabbers, some digitize the image line by line (line grabbers), so having a full frame digitized can take as long as 30 seconds. On the other hand, frame grabbers are real-time digitizers, that is, there is no apparent delay between the input of the video signal and its digitization. They digitize 25 to 30 frames per second. Frame grabbers have their own memory, typically receive a composite video signal and can produce a RGB or composite output video signal.

Video to VGA boards are digitizers that depend on your VGA card to display the image on your computer monitor. They are cheaper than frame grabbers, and can produce reasonable quality monochrome, 256 levels of gray or color images (but see below on resolution). Some can operate at frame grabber speed. If you want to digitize color you should keep two things clear: one thing is the ability to digitize color in RGB and store it in a file and quite another is the ability of your PC to display the 16 million colors potentially available in

a RGB file. This later capability depends mainly on the video memory available for display. 512 K of video memory can display a selection of 256 colors, and 1 Megabyte up to 32,768 colors. Humans can distinguish up to 350,000 colors (Lindley, 1991).

The main points to consider in selecting a digitizing board are:

1) "Resolution." This is usually reported in pixels such as 320x200 or 640x480. Though the image may be divided into this number of pixels, the resolution is best thought of in terms of vertical and horizontal lines as discussed earlier.
2) Real time versus time-delay digitizing. Real time gives you better control for digitizing what you want, but for non-moving objects some delay can be tolerated.
3) Different kinds of video input, composite, S-Video, RGB, etc., and output signal.
4) Very important is the availability of ready to use software for the application you have in mind. Also, it is interesting to consider boards with some operations (image compression, image treatment, etc.) implemented in hardware. Of the morphometric software tested later, MorphoSys, Java (and Coda) will not run at all, unless you have installed the proper frame grabber in the computer. MTV does not need the presence of a frame grabber to run (although you will need it if you want to acquire images with it or work on the TV and PC monitor at the same time). At the time of writing this, MorphoSys and MTV are limited to working with Imaging Technology PCVISIONplus and VISIONplus AT-OFG frame grabbers. JAVA is able to work with PCVISIONplus, Truevison TARGA M8, Metrabyte MV-1 and Data Translation 2855, 2953. So, the only software that can really work with any video digitizing board is MTV, first, because it does not need the presence of any frame grabber to run; second, because it can work on the PC monitor with any image saved in TIFF format.

We have worked with both Imaging Technology boards,PC Vision Plus and Vision Plus AT OFG that only accept composite video input and two Video-VGA boards, LifeView that accepts composite video input and Screen Machine that accepts both composite and S-Video signals. These Video-VGA cards could work with PAL or NTSC.

Personal Computers (PCs)

Since a computer is a very common piece of equipment in everybody's lab, we will only reiterate the minimum configuration for an image system. You should have a VGA card with adequate memory for color (512K for 256 colors). The faster your processor (386 or 486) and the larger your hard disk the better. You must have at least one free slot to insert the digitizing card (but see below).

A 486DX has a built in math coprocessor. The latter is a low priced accessory for 386 and 286 computers. Math coprocessors are highly recommended for digitizing, compression (discussed later), and a necessity for morphometric analysis.

Another factor to take into account is that there are three standards of expansion buses: Industry Standard Architecture (ISA), Extended Industry Standard Architecture (EISA), and the MicroChannel Architecture (MCA). This is important because the video digitizing cards you acquire must go in a slot (as mentioned before) and be compatible with the standard your computer belongs to.

The ISA standard corresponds to the old designation, AT bus. The EISA standard is compatible with the ISA standard. The MCA is totally incompatible with the other two standards.

Monitors

We comment here on both the computer monitor and the TV monitor if used. Some monitors with multiscan capability (e.g., the Sony GVM-1400 QM) are able to display video and CGA/EGA/VGA signals. With this monitor you can switch sequentially between the two modes. This is not a comfortable solution because for the same price you can have two separate monitors, one for each function.

Some characteristics you should look at when comparing computer or TV monitors are: dot pitch size, horizontal/vertical resolution and sychronization range. We have used a PC NEC Multisync 2A, two Sony TV monitors: GVM 1400QM and PVM 1342Q; and a Toshiba green display with composite input, PA 7150E.

Additional hardware

Some additional equipment can provide images to your programs, that would not be possible in any other way . An alternative to a VCR for temporary storage of a video frame can be a device like the Sony XV-D300 Digital Video Adaptor, where you can leave a single frame in analog format and then send it to a grabber or other destination.

To properly digitize recorded images frame by frame a TBC (time basis corrector) is necessary. It is true that most digitizing programs can freeze an image as it passes through the frame grabber, but this is not the same as digitizing a frozen image in the camera or VCR. When the video signal that goes from a tape to the frame grabber is frozen by it, there is normally no time basis error and the image is frozen without distortion (jitter). If you stop an image on your tape and input it to the frame grabber, you will probably have distortions due to the

fact that the mechanical motion of the recorder head drum is not smooth enough to reproduce a stable frame (see, for example, Blinn, 1990 or Luther, 1991, for an extended explanation). These time base errors are corrected by TBCs. TBCs used to be an expensive piece of equipment for the video professional (in Spain, at least, TBCs for video professionals run between $5,000 to $10,000). Fortunately, the everyday more demanding amateur video market is providing this equipment at more reasonable prices (the Panasonic WJ-AVE5, that includes a TBC, tested by us runs around $1500). These machines not only correct for time base errors but also memorize one or several frames (some can display the two fields of a frame one at a time) so they can be used as independent video sources, with video input from a video device, or a transformed image file. Some new amateur cameras and some video digitizers include TBC functions.

Image storage

Images require large amounts of memory. One 640x480 pixel image in 24 bit color (8 bits for each color) requires about 900 Kbytes of memory, or 300 Kbytes with 256 shades of gray or 256 colors. Very few images may be stored on a floppy disk (only one with 24 bit color), and hard disks become filled quickly. Magneto-optical and laser disks of 600 Mbytes capacity are available, but are expensive at this time. They are also slower than magnetic disks (access time to some hard disk below 20 ms, and around 70 ms in magneto-optical drives), but their access time may soon compete with magnetic disks. Other hardware to look at are Compact Disk Technology and their analog counterpart, Laser Disks. A review of these components is beyond the scope of this paper.

SOFTWARE

Two kinds of software are worth considering:
a) Software for image and data acquisition, and possibly image enhancement.
b) Software for format transformation and image compression.

Software for data and image acquisition

To our knowledge there are only three packages specifically written for morphometric data acquisition, although other software include measurement features as part of their capabilities. These are MorphoSys, MTV and Coda. The later two are designed for coordinate data only. MorphoSys can work with coordinates and automatically generate outlines; macros can be written for quite sophisticated acquisition.

We include a comparison of the accuracy of MTV and Morphosys (CODA was unavailable). We also have examined JAVA, a program with more features, as it includes apart from morphometric measurement, densitometry, object counting, and image enhancement. Java costs around $2000 while MTV and MorphoSys cost $350 and $250 respectively. These last two are very inexpensive for the specialized routines they offer, and they make working with morphometric data very simple.

The great advantage of MTV over MorphoSys and JAVA is that MTV can work on imported images (TIFF files) without using a frame grabber (just having an EGA or VGA card) directly on the computer screen (JAVA works with TIFF files, but you still need to have the frame grabber in the computer). In this way, one institution can have a main image acquisition system to take and write images files with a single frame grabber, and as many secondary working units as desirable without the expense of additional frame grabbers. At present, MorphoSys can import/export image files only for the PCVISIONplus grabber.

Software for image transformation and compression

Images may be saved in files using several possible formats. There are many graphics formats available and it is not possible here to make even a small summary (see Rimmer, 1990 for an introduction). It is useful to convert images from one format to another, to make images available to other software, and to make other images available to your software. We have used a very flexible program called HIJAAK for format conversion. PIZAZZ PLUS and HOTSHOT GRAPHICS can do many of the same conversions. Some shareware programs like GWS (see appendix) include some limited conversion routines. Image file transformation can provide morphometric programs with images taken from a scanner, a still video camera or other image input device. Images may also be moved between MacIntosh and IBM PC's using some of this software.

MorphoSys stores image memory from the board to an *.IMG file (incidentally this has nothing to do with IMG files from Digital Research). MTV writes a slightly different format called ITI (ITI is the file format name used by IMAGING TECHNOLOGY, the PC VISION manufacturer) and also TIFF files. The latter is a common image file format in computer graphics. JAVA also, saves images in TIFF and other formats. We have included in the diskette accompanying this book an utility for image file format conversion between MorphoSys and MTV. This utility can translate TIFF files generated by MTV to IMG format files (used by MorphoSys) and viceversa.

Image files usually take a lot of memory space. ITI and TIFF formats usually take (uncompressed) around 300 kbytes of memory, so unless you have a big hard disk space you could run out of space very soon. Despite the massive storage devices we have mentioned earlier (the magneto-optical disks), there are some

alternative ways to save memory space with image files, and a brief comment on image compression follows.

There is some difference between compression of color and black and white images, and four parameters are considered in compression techniques (Wayner 1991):

1) the compression ratio or number of times the initial size of the image file is reduced, for example 10 to 1;
2) speed of the compression algorithm;
3) Loss of data in the compression/decompression cycle - Some algorithms are "lossless" because in one cycle no data is lost, while others are called "lossy" because they exchange data loss for improved compression ratios;
4) image quality, although this is not an objective criterion.

There are several compression standards, and some give quite good compression ratios. One of them is the JPEG (Joint Photographic Expert Group) standard. It is important to check if the frame grabber you acquire has JPEG file compression capability implemented in hardware. This cuts down the compression/decompression cycle time.

TESTS ON RESOLUTION, AND MEASUREMENT PRECISION AND ACCURACY

Three things were tested with the hardware and software available to us: resolution, focal length and working distance selection and software accuracy. We suggest that, at least, resolution should be tested with the equipment available to you. For testing resolution of our components we used a standard TV chart (Fig.2). It is easy to use one to see how many horizontal and vertical lines the system is able to discriminate. This can be done first with the camera connected directly to the monitor, and then again by connecting the camera to the monitor via the frame grabber, to observe the decrease in resolution. Loss of resolution due to taping can be detected looking at the TV chart once taped. The tables offer the following checks of resolution: a) video cameras to monitors, by sending a video signal before or after taping; b) video cameras to frame grabbers or Video-VGA boards and then to TV or PC monitors (Table I).

It is surprising that the Toshiba PA7150 E monitor, a green computer monitor that came with a pre–PC computer (the toshiba T–100, a 12 years–old–computer) that accepts composite signals, could give a better resolution than high–priced Sony monitors. Neither Sony nor Canon Spanish technical staff gave us a satisfactory explanation, and as we are not engineers, we prefer not to guess. But it opens up the possibility of using low–priced green monitors with composite input, at least for working with this kind of video output. ITI frame grabbers, when tested with high resolution cameras (in this case, only the Cohu B&W) gave the highest resolution on display. The Sony Hi8 V–5000E composite output produce a vertical undulating movement through frame grabbers and

SOURCE	TYPE OF SIGNAL	MONITOR	CAMERA			
			Sony F-555E (PAL)	Sony V-5000E (PAL)	Canon EX-1 (PAL)	Cohu (NTSC)
Live	Composite	Sony	280	300	300	420
		Toshiba	320	300	380	520
	S-Video	Sony	—	380	380	—
Recorded (8 mm)	Composite	Sony	< 200	< 200	< 200	—
		Toshiba	< 200	< 200	< 200	—
	S-Video	Sony	—	—	—	—
Recorded (Hi8)	Composite	Sony	—	280	280	—
		Toshiba	—	320	320	—
	S-Video	Sony	—	320	320	—

DIGITIZING CARD	TYPE OF SIGNAL	MONITOR	CAMERA			
			Sony F-555E (PAL)	Sony V-5000E (PAL)	Canon EX-1 (PAL)	Cohu (NTSC)
Life-Video	Composite	Computer	260	300	300	250
Screen Machine	Composite	Computer	280	290	240	—*
	S-Video		—	280	380	—
AT-OFG	Composite	Sony	320	320	340	—
		Toshiba	340	320	360	—
PCVision Plus	Composite	Sony**	—	—	—	450

* The signal from the camera is not clear.
** Using MorphoSys, with a Sony PVM 1342 Q monitor.

Table I

Horizontal resolution (in lines) for:

a) Four cameras, depending on the source being live or recorded video.
b) Four digitizing cards, using four different cameras.
The cameras are:
— Sony F-555-E, 8 mm camcorder (PAL).
— Sony V-5000-E, Hi8 camcorder (PAL).
— Canon EX1, Hi8 camcorder (PAL).
— Cohu 4815–2000, video camera (NTSC).
The monitors are:
— Sony PVM 1342 Q.
— Sony GVM 1400 QM.
— Toshiba PA 715 OE.
The digitizing cards are:
— AT–OFG from Imaging Technologies, Inc.
— PCVision Plus from Imaging Technologies, Inc.
— Life View Video Board from Animation Technologies, Inc.
— Screen Machine from Fast Electronics GmbH.

Video–VGA boards, resulting in poorer resolution than the Canon EX1. The NTSC option of the Screen Machine Video–VGA board did not work properly with the Cohu.

Focal length and working distance selection

A possible way to select an appropiate focal length and working distance for a specific object is: size of the object and size of the image in the focus plane of the optical system are related in the following way (Blaker, 1976):

$$\frac{h'}{F} = \frac{h}{v}$$

where h' is the image size in the focal plane
F is the lens focal length
h is the object size, and
v is the lens–to–subject distance

The image you see in the monitor is «h» multiplied by a constant. The constant can be estimated by measuring the image on the screen.

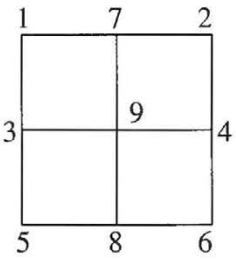
We kept the image size on the screen constant and filling slightly more than half of the screen. We then adjusted the other three variables in order to ascertain which combination is best for every object we measured. In the example given we tried to measure inaccuracies by working at three different focal lengths: 25 mm, 60 mm and 120 mm; and two distances from the lens to the object in focus. We varied the size of the object by using various portions of a piece of millimeter paper.

Results are reported in Table II. The example worked here was done with the Canon EX1, with the 8–120 mm objective and the software MorphoSys.

Focal Length (mm)	Dist. Obj. (cm)	Mean Absolute Error and Standard Deviation				
		Diagonal (mm)	Horizontal (mm)	Vertical (mm)	Total area (mm^2)	Partial area (mm^2)
25	70	MAE: 2.21 SD: 0.14	MAE: 1.14 SD: 0.31	MAE: 1.78 SD: 0.20	MAE: 252.1 SD: 7.87	MAE: 56.6 SD: 8.86
	142	—	—	—	—	—
60	70	MAE: 0.96 SD: 0.14	MAE: 0.75 SD: 0.081	MAE: 0.55 SD: 0.06	MAE: 55.00 SD: 1.74	MAE: 12.83 SD: 1.15
	142	MAE: 3.68 SD: 0.14	MAE: 2.54 SD: 0.27	MAE: 2.36 SD: 0.24	MAE: 423.4 SD: 8.16	MAE: 96.2 SD: 5.69
120	70	MAE: 0.36 SD: 0.05	MAE: 0.22 SD: 0.05	MAE: 0.23 SD: 0.07	MAE: 10.19 SD: 0.77	MAE: 2.13 SD: 0.72
	142	MAE: 0.84 SD: 0.07	MAE: 0.38 SD: 0.13	MAE: 0.68 SD: 0.10	MAE: 48.1 SD: 3.40	MAE: 10.05 SD: 1.70

Table II

For each case we have taken measurements with millimeter paper (see below). Measurements at focal length 25 mm and 142 cm from object could not be taken properly. The measurements have been taken on the following square:



The measures were:

Diagonals: 1-6, 2-5
Horizontals: 1-2, 3-4, 5-6
Verticals: 1-5, 7-8, 2-6
Total area: 1-2-6-5
Partial areas: 1-7-9-3, 7-2-4-9, 9-4-6-8, 9-8-5-3

The number of replications were:

Diagonal measures: 20;
Horizontal measures: 30;
Vertical measures: 30;
Total Area measures: 20;
Partial Area measures: 40;

The real values were:

Focal length: 25 mm, 60 mm; Dist. Obj.: 70 cm, 142 cm.
Diagonal: 113.14 mm
Horizontal: 80 mm
Vertical: 80 mm
Total Area: 6400 mm^2
Partial Area: 1600 mm^2

Focal length: 60 mm, 120 mm; Dist. Obj.: 70 cm, 142 cm.
Diagonal: 56.57 mm
Horizontal: 40 mm
Vertical: 40 mm
Total Area: 1600 mm^2
Partial Area: 400 mm^2
Focal length: 120 mm; Dist. Obj.: 70 cm.
Diagonal: 28.28 mm
Horizontal: 20 mm
Vertical: 20 mm
Total Area: 400 mm^2
Partial Area: 100 mm^2

MAE = Mean Absolute Error
SD = Standard Deviation.

Looking at the table, it can be seen that the best focal length in terms of accuracy is 120 mm at 70 cm. from object. Keeping focal length constant and changing object size (and distance to object if screen image is constant too) can be done to test the adequacy of a single focal length for working properly with different lenses.

Software and accuracy

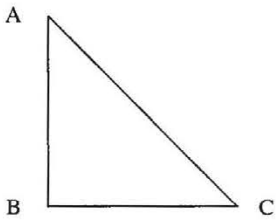
Accuracy of measurements of MORPHOSYS, MTV and JAVA has been tested in a PC VISION PLUS 640 NTSC frame grabber with a COHU 4815 Series Monochrome camera. We could not make the same test with the AT–OFG board because MTV

Distance to the object 113-114 cm Focal lenght 75 mm		Scale over the X axis*			Scale over the Y axis*		
		1	2	3	1	2	3
Distance AB	MorphoSys	1.22 (0.39)	0.76 (0.18)	0.36 (0.11)	—	—	—
	MTV	-0.18 (0.33)	-0.52 (0.08)	-0.37 (0.00)	0.88 (0.06)	0.25 (0.09)	0.00 (0.00)
	Java	-0.78 (0.32)	-0.17 (0.06)	-0.02 (0.13)	0.75 (0.35)	0.06 (0.20	0.02 (0.06
Distance BC	MorphoSys	1.20 (0.34)	1.11 (0.27)	0.42 (0.15)	—	—	—
	MTV	0.44 (0.23)	-0.02 (0.11)	-0.19 (0.00)	1.75 (0.13)	0.68 (0.16)	0.30 (0.10)
	Java	-0.14 (0.28)	0.04 (0.20)	0.04 (0.12)	1.39 (0.43)	0.36 (0.19)	0.17 (0.06)
Distance CA	MorphoSys	2.04 (0.23)	1.47 (0.18)	0.68 (0.09)	—	—	—
	MTV	-0.13 (0.11)	-0.48 (0.09)	-0.52 (0.13)	1.80 (0.31)	0.69 (0.08)	0.15 (0.17)
	Java	-0.41 (0.43)	-0.02 (0.06)	-0.04 (0.14)	0.55 (0.65)	0.59 (0.09)	0.10 (0.13)
Area	MorphoSys	87.9 (8.8)	47.13 (4.44)	6.91 (1.91)	—	—	—
	MTV	5.2 (15.1)	-4.86 (6.36)	-4.85 (1.78)	93.5 (24.7)	18.36 (3.03)	1.14 (1.33)
	Java	60.6 (31.3)	18.16 (12.45)	3.45 (2.70)	87.6 (21.9)	21.94 (15.53)	6.69 (3.47)
Angle	MorphoSys	0.12 (0.10)	-0.05 (0.11)	-0.003 (0.44)	—	—	—
	MTV	0.00 (0.00)	0.02 (0.07)	0.00 (0.00)	-0.02 (0.05)	0.05 (0.17)	0.00 (0.00)
	Java	-0.06 (0.07)	0.14 (0.14)	0.00 (0.00)	0.08 (0.08)	0.22 (0.12)	0.05 (0.17)
Angle CAB	MorphoSys	-0.09 (0.19)	0.28 (0.20)	-0.01 (0.31)	—	—	—
	MTV	0.38 (0.15)	0.40 (0.16)	0.48 (0.17)	0.45 (0.02)	0.52 (0.13)	0.85 (0.25)
	Java	0.33 (0.11)	0.28 (0.10)	0.24 (0.24)	0.24 (0.08)	0.31 (0.07)	0.16 (0.14)
Angle BCA	MorphoSys	-0.07 (0.22)	-0.47 (0.22)	0.09 (0.26)	—	—	—
	MTV	-0.35 (0.07)	-0.38 (0.15)	-0.28 (0.19)	-0.38 (0.07)	-0.28 (0.10)	-0.32 (0.17)
	Java	-0.32 (0.12)	-0.39 (0.14)	-0.37 (0.19)	-0.23 (0.06)	-0.37 (0.12)	-0.39 (0.14)

Table III

Table of mean absolute error and standard deviation (between parentheses).

The measures have been taken on the following triangle. Number of replications= 10.



1: The triangle fills the whole display.
Real values:
Distance AB: 70 mm Angle ABC: 90°
Distance BC: 70 mm Angle CAB: 45°
Distance CA: 98.995 mm Angle BCA: 45°
Area: 2450 mm^2

1: The triangle fills nearly half the display:
Real values:
Distance AB: 40 mm Angle ABC: 90°
Distance BC: 40 mm Angle CAB: 45°
Distance CA: 56.5685 mm Angle BCA: 45°
Area: 800 mm^2

1: The triangle fills nearly a quarter of the display:
Real values:
Distance AB: 20 mm Angle ABC: 90°
Distance BC: 20 mm Angle CAB: 45°
Distance CA: 28.28 mm Angle BCA: 45°
Area: 200 mm^2

did not work properly on it. Several errors detected by us in the version for the AT–OFG board (European version) are being corrected (Garr Updegraff, pers. comm.). See Table III for a explanation of the test and a summary of the results.

It can be seen that, the three programs are very accurate, MTV and JAVA performance being slightly better than MORPHOSYS.

CONCLUSION

It is difficult to make satisfactory recomendations, but some points deserve attention when looking at equipment for morphometric analysis:

a) Resolution is a good variable that can be used to check the perfomance of the equipment.

b) Software independence from grabber (at least for some operations) is a desirable feature.

c) The Video--VGa seems to be a promising kind of board, but it is not as good as the true frame grabbers.

d) If you assemble the different compenents of your equipment, be sure that all inputs/outputs are compatible empirically.

e) Do not pay more for something that is «theoretically» better, unless you verify it is true.

ACKNOWLEDGEMENTS

This chapter has greatly benefited with the comments of L. Marcus. J. Rohlf, F. Bookstein, G. Updegraff, B. Weinstein and L.M. Carrascal also gave useful comments. C. Acedo (Fonovideo) and Rogelio Sanchez (MNCN) helped with hardware and technical questions. DGICYT PB89–0044.

APPENDIX

Addrex of Hardware/Software mentioned in the article:

DYCAM

Dycam, Inc., 9588 Topanga Canyon Blvd., Chatsworth, CA 91311 USA.

EXETER

Exeter Software, 100North Country Rd., Bldg. B. Setauket, N.Y. 11733, USA. (Supplies MTV, MorphoSys and other scientific software).

FOTOMAN

Logitech, Inc. Fremont, CA 94555, USA

GRAPHICS WORKSHOP (GWS). This program can be retrieved by anonymous FTP from OAK.OAKLAND.EDU, in the directory pub/msdos/graphics.

HIJAAK

Inset Systems, 71 Commerce Drive, Brookfield, CT 06804, USA.

HOTSHOT GRAPHICS

SymSoft, 444 First Street, Los Altos, CA 94022. USA

IMAGING TECHNOLOGY

Imaging Technology Inc., 600 West Cummings Park, Woburn, MA 01801-6343, USA.

ION (XAXHOP IN USA)

Canon Inc. 2-7-1 Nishi-Shinjuku-ku, Tokyo 163, Japan

JAVA

Jandel Scienfic, 600 West Cummings Park, Woburn, MA 01801-6343, USA

MORPHOSYS

Lynch, Marks&Associates. 2180 Dwight Way#C Berkeley CA 94704 USA (Whole system, hardware and software).

PIZAZZ PLUS

Application Techniques, Inc. 10 Lomar Park Drive. Pepperell, MA 01463, USA

REFERENCES

ANONYMUS, 1992. Data Sources. Ziff-Davis Publication, N.Y.

BEALE, S., CAVUOTO, J., 1991. *The Scanner Handbook*.. Heinemann Newtech, Oxford. (In USA: Micro Publising Press).

BLAKER, A. A., 1976. *Field Photography*. W. H. Freeman and Company.

BLINN, J. F., 1990. Wonderful World of Video. IEEE *Computer Graphics & Applications*, May: 83-87.

FINK, W. L., 1987. Video Digitizer: A System for Systematic Biologists. *Curator* 30/1: 63-72.

FINK, W. L., 1990. Data Acquisition for Mophometric Analysis in Systematic Biology. Pp. 9-19. In: Rohlf, F. J. & Bookstein, F. L. (eds.) *Proceedings of the Michingan Morphometrics Workshop*, Museum of Zoology special publication no.2, University of Michigan: Ann Arbor.

GLOVER, F., 1990. *Image scanning for desktop publishers*. Windcrest.

JÄHNE, B., 1991. *Digital Image Processing. Concepts, Algorithms and Scientific Applications*. Springer-Verlag, Berlin.

LINDLEY, C. A., 1991. *Practical image processing in C. Acquisition, manipulation, storage*. John Wiley&Sons, Inc. New York.

LUTHER, A. C., 1991. *Digital video in the PC environment*. 2nd edition. McGraw-Hill, New York.

MACLEOD, N., 1990. Digital Images and Automated Image Analysis Systems. Pp. 21-35. In: Rohlf, F. J. & Bookstein, F. L. (eds.) *Proceedings of the Michigan Morphometrics Workshop*. Museum of Zoology special publication no.2, University of Michigan: Ann Arbor.

MARCUS, L. F., 1991. Computer Programs for Multidimensional Palaeobiology. Pp. S1-S39. In: Reyment, R. A. *Multidimensional Palaeobiology*. Pergamon Press.

MEACHAM, C. E., 1992. Practical Machine Vision for Morphometric Data Acquisition. In: Sorensen, J. T.&R. G. Footitt, eds. *Ordinations in the Study of Morphology, Evolution, and Systematics of Insects: Applications and Quantitative Genetic Rationales*. Elsevier. Amsterdam.

RIMMER, S., 1990. *Bit-Mapped Graphics*. Windcrest Books.

ROHLF, F. J., 1990a. An Overview of Image Processing and Analysis Techniques for Morphometrics. Pp. 37-60. In: Rohlf, F. J. & Bookstein, F. L. (eds.) *Proceedings of the Michigan Morphometrics Workshop*. Museum of Zoology special publication no.2, University of Michigan: Ann Arbor.

ROHLF, F. J., 1990b. Morphometrics. *Annual Review of Ecology and Systematics*. 21:229-316.

SOKAL, R. R. & ROHLF, F. J., 1981. *Biometry*. W. H. Freeman and Company.

WAYNER, P., 1991. Inside Quicktime. *Byte* 16(13):189-196.

PART THREE

METHODOLOGY AND SOFTWARE

SOME ASPECTS OF MULTIVARIATE STATISTICS FOR MORPHOMETRICS

LESLIE F. MARCUS

Department of Biology
Queens College of CUNY
Flushing, NY 11367

Department of Invertebrates
American Museum of Natural History
CPW at 79th
New York, NY 10024
LAMQC@CUNYVM.BITNET, LAMQC@CUNYVM.CUNY.EDU

CONTENTS

ABSTRACT

Some multivariate statistics procedures widely used for distance and coordinate data are explained with data examples. An extensive discussion of Gabriel's biplot method is provided together with a computer program to produce several forms of the biplot.

The relation between univariate t tests and Hotelling's T^2 test statistic is given in an intuitive way, and similarly the relation between univariate and multivariate analysis of variance is explored.

INTRODUCTION

Traditionally, the description of form - the size and shape of organisms, has been based on linear measurements or distances between pairs of reference points. These distances have been the raw data for multivariate descriptive and inferential statistics in systematics, ontogeny and functional analysis. Distances are frequently transformed using logarithms, and a favorite form of multivariate analysis has been principal components.

The new morphometrics, advocates that size and shape comparisons are best captured in organism space recorded by two or three dimensional coordinates of homologous landmarks. In the twentieth century, D'Arcy Thompson provided graphical methods for displaying change and deformation in shape. Bookstein (1978) developed methods for quantifying D'Arcy Thompson type shape transformations, and more recently (1991) has provided methods for analyzing the coordinates themselves. The "truss" formed a kind of bridge between the two perspectives, where distances could be converted to coordinates, when the organism was covered by an adequate number of distance measures described in the form of sufficient triangles.

Landmark data may be transformed to Bookstein shape coordinates (earlier called shape coordinates) and centroid size sequestered. Summary statistics then describe size and for the new coordinates, uniform or linear, e.g. affine, shape differences and non-linear shape differences (Bookstein, 1990,1991). The landmarks as shape coordinates, and other derived size and shape statistics may then be analyzed using classical multivariate statistics (see Rohlf, this volume). Rectangular data matrices of raw or derived statistics form the bases for analyses both for the more traditional distance, and the coordinate based methods.

We will consider here that each row of the data array contains all of the values measurements, derived coordinate based values for a single specimen or other operational unit (see first array below). The columns will contain all of the measurements of one character, variable, coordinate or feature for all specimens, taxa or other kinds of units. Additional columns, or auxiliary matrices, may be used to give specimen number, locality, sex, environmental features and other descriptive data which may be used to partition our data into subsets; or for further

analysis correlating the measurements with extrinsic and intrinsic non-morphometric variables.

For both distance and coordinate derived data, frequency plots or histograms summarize information for single variables, while bivariate plots or scattergrams visually summarize patterns of variation two variables at a time. I strongly recommend lots of such plots to "get a feel for the data" and as a method for discovering outliers and errors in measurement or data identification. Note that because coordinate data are in the space of the organism then two dimensional or three dimensional projection plots of all landmarks simultaneously, or overlays on two dimensional outlines can be useful also. They usually require some kind of Procrustes superimposition free of the affects of the original digitizing orientation and location, and sometimes adjustment for size. The GRF program of Rohlf (1990) provides such a tool for overlays in two dimensions (Rohlf, 1990 and Rohlf & Slice, 1990).

Scatter in many dimensions extends scattergrams, but is impossible to depict past two dimensions and even hard to interpret in two dimensional projections of three dimensions. NTSYS (Rohlf, 1990) and NCSS (Heintze, 1990) for example give us tools for simultaneously examining a number of bivariate plots for the same data array. Both have routines for spinning three dimensional scatters to study and examine data points three variables at a time.

Summaries of a data array in the form of means, variances and covariances and correlations complete the basic descriptive statistics of classical univariate and multivariate statistics. These can be used to produce graphic projections of the data for further statistical analysis and as a first step in multivariate inference.

The most popular method for summarizing multivariate scatter, among linear metric methods, has been principal components analysis, a method which provide displays in the lowest possible dimension summarizing the maximum variance and covariance for multidimensional data. A corollary to the variance summarizing feature, is that the between specimen distances in the space of the variables are summarized best, in a sum of squares sense, as well. This fact is exploited in Gower's principal coordinate analysis (Marcus, 1990).

Tests of hypothesis on differences in sexes, ontogenetic stages, localities or taxa extend our usual univariate t and Analysis of Variance procedures to Hotelling T^2 and Multivariate Analysis of Variance (MANOVA) for testing hypotheses about joint means or centroids. Distance statistics such as Mahalanobis D^2 or D, are closely related to T^2 as will be shown below. D is a useful unit free descriptive statistic for comparing multivariate centroids and for identification using discriminant analysis. Canonical variate ordinations provides a way of displaying centroids and scatter in a reduced dimensional space.

All of these topics are well described in the literature (Marcus, 1990; Reyment, 1990; Reyment, 1991; Dillon & Goldstein, 1984; Seber, 1985;

Krzanowski, 1988; and many others). I still see so much uncertainty in there use, that illustrations with biological data should help some researchers. It is not always easy to run computer programs in the various packages so in an appendix I will provide directions for using MATLAB procedures for manipulating all of the data sets discussed in the text. These programs and all of the data sets described are available on the disk accompanying this book.

THE SINGULAR VALUE DECOMPOSITION AND THE BIPLOT

Principal components analyses are usually based on the variance-covariance or correlation matrix among variables. The eigenvectors and eigenvalues of either matrix give us a unique decomposition of variance and covariance into principal components. The only circumstance where the decomposition is not unique is when two eigenvalues are exactly the same, but this rarely occurs for real data. A small set of principal component scores, the number of principal components depending on the magnitude of eigenvalues or some other criterion, summarize our data best in a least squares sense.

The singular value decomposition provides an alternative computation technique for doing a principal component analysis giving the same results, but it is based on a mean centered or column standardized data matrix. The covariance or correlation matrix need not be part of the analysis, though both may form useful adjuncts for understanding and interpretation.

The singular value decomposition (svd) is a direct dissection of the data matrix, usually mean centered, or any matrix into the product of three matrices. The better known terms for the results of a principal components analysis based on a covariance or correlation matrix (eigenvalues, eigenvectors and PC scores) are given in parentheses for those not familiar with the svd. The singular value decomposition describes the data array in terms of scores (scaled PC scores), the singular values (simple functions of the eigenvalues), and loadings (eigenvectors). Here is some data, the first 10 specimens and 5 variables used in the *Zygodontomys* data example in Marcus (1990) and Marcus and Corti (1989). These are measurements of small rodent skulls from one locality in South America. We will only consider for now the first two variables, condylo-incisive length (CIL) and length of the diastema (LD). I have demarcated the relevant parts of the data (Note all results are directly extracted from the MATLAB program BIPLOT4.M supplied on the accompanying disk with documentation for running). Measurements are in millimeters and are recorded to the nearest 0.05 mm. using hand held dial calipers (data courtesy of R. Voss; see Voss *et al.*, 1990 and Marcus & Corti, 1989 for details).

CIL	LD	LM	BM1	LIF
23.00	6.20	4.05	1.30	5.05
24.20	6.45	4.30	1.25	5.25
24.90	6.95	3.90	1.15	5.70
24.65	6.75	4.10	1.20	5.65
24.75	7.00	4.00	1.20	5.30
25.50	7.15	4.10	1.20	5.60
25.30	7.35	4.05	1.30	5.85
24.75	6.85	4.10	1.20	5.55
25.05	6.85	4.35	1.35	5.50
25.15	6.85	4.10	1.25	5.65

The means and standard deviations are:

	CIL	LD	LM	BM1	LIF
Mean	24.725	6.840	4.105	1.240	5.510
Std. Dev.	0.707	0.327	0.132	0.061	0.241

The data as deviations from the mean are:

CIL	LD	LM	BM1	LIF
−1.725	−0.640	−0.055	0.060	−0.460
−0.525	−0.390	0.195	0.010	−0.260
0.175	0.110	−0.205	−0.090	0.190
−0.075	−0.090	−0.005	−0.040	0.140
0.025	0.160	−0.105	−0.040	−0.210
0.775	0.310	−0.005	−0.040	0.090
0.575	0.510	−0.055	0.060	0.340
0.025	0.010	−0.005	−0.040	0.040
0.325	0.010	0.245	0.110	−0.010
0.425	0.010	−0.005	0.010	0.140

The variance covariance matrix is:

	CIL	LD	LM	BM1	LIF
CIL	0.5007	0.2089	−0.0004	−0.0089	0.1411
LD	0.2089	0.1071	−0.0119	−0.0040	0.0637
LM	−0.0004	−0.0119	0.0175	0.0051	−0.0073
BM1	−0.0089	−0.0040	0.0051	0.0038	−0.0032
LIF	0.1411	0.0637	−0.0073	−0.0032	0.0582

and correlation matrix:

	CIL	LD	LM	BM1	LIF
CIL	1.0000	0.9020	0.0045	0.2044	0.8265
LD	0.9020	1.0000	0.2748	0.1988	0.8069
LM	0.0045	0.2748	1.0000	0.6223	0.2282
BM1	0.2044	0.1988	0.6223	1.0000	0.2173
LIF	0.8265	0.8069	0.2282	0.2173	1.0000

For now, we need only consider the first two means, and upper left 2 by 2 parts of the last two matrices above. Later we will analyze all five variables, and then the complete data set of 68 specimens and 12 variables from which this data was extracted.

The eigenvalues and eigenvectors of the 2x2 covariance matrix are:

	L=Eigenvalues		Eigenvectors (also see U below)		
	Value	cum. %		PC1	PC2
PC1	0.5909	97.2169	CIL	0.9181	–0.3964
PC2	0.0169	100.0000	LD	0.3964	0.9181

Note that the sum of variances 0.6078 is the same as the sum of the eigenvalues, as is always true for principal component analysis.

The three parts of the singular value decomposition are:

V

–0.7968	0.2467
–0.2760	–0.3843
0.0886	0.0810
–0.0453	–0.1356
0.0375	0.3511
0.3618	–0.0580
0.3166	0.6158
0.0117	–0.0019
0.1311	–0.3067
0.1709	–0.4083

$D=((n-1)L)^{0.5}$

2.3061	0
0	0.3902

	U	
CIL	0.9181	–0.3964
LD	0.3964	0.9181

Popularly scaled PC scores having variances equal to the eigenvalues are obtained from post multiplying V by D.

VD=Usual	PC scores
–1.8374	0.0962
–0.6366	–0.1499
0.2043	0.0316
–0.1045	–0.0529
0.0864	0.1370
0.8344	–0.0226
0.7301	0.2403
0.0269	–0.0007
0.3023	–0.1197
0.3941	–0.1593

DV'VD/(n-1) = L to show that these values of VD produce the eigenvalues or variances of our usual PC's on the diagonal below.

Note their sum is again equal to the sum of the variances.

0.5909	0.0000
0.0000	0.0169

Here we give also VDU' to complete the process and show that the original deviations from the mean are recovered from the product of the three matrices. If we add back the means of course we recover the original data.

CIL	LD
–1.725	–0.640
–0.525	–0.390
0.175	0.110
–0.075	–0.090
0.025	0.160
0.775	0.310
0.575	0.510
0.025	0.010
0.325	0.010
0.425	0.010

For more than two variables, if two principal components adequately summarize our data matrix, then a biplot is a an adequate display of both the specimen scores and the variable "scores" on the same graph, and most of the

original data can be recovered. Though Gabriel suggested the method in 1968 it has been little used. There has been a recent resurgence in interest in the biplot graph for supplementing a principal components analysis as the singular value decomposition has become more widely available in computer packages (eg. SAS PROC IML, MATLAB, and NTSYS). The biplot does not have to use the singular value decomposition and can be derived from an ordinary PC analysis.

Its use is advocated here provided enough information is summarized in two dimensions. In any case the biplot will be the best two dimensional display of our data in an over all sums of squares sense, but interpretation of relations between points, and variable vectors will only be appropriate to the degree that two principal components summarize the data. Plots of only the first two principal components have all too often been published, when they do not adequately summarize the data and an inadequate analysis of residuals is done. Jackson (1991 book on Principal Components) devotes quite a bit a material to this most important area and Marcus (1990) gives some brief suggestions.

An attempt at illustrating the biplot was given in Marcus (1990), where PC scores and variable vectors were presented on separate graphs and not superimposed. It has been common to plot the specimens as points, and variables as vectors on one graph -following an example in Gabriel (1968). Gabriel found it necessary to use separate scales for the specimens and variables. Here a "fudge" factor is introduced in the software, to provide more pleasing graphs (see software instructions in the Appendix). It is a convention and a useful one as we will see below. In Figure 1 for the same data the plots are superimposed. Three forms of the biplot are given as discussed in detail below, and in addition a fourth plot which is not a true biplot but has some of its features. Other data examples will be given.

The first biplot figure 1a is an exact biplot as the two variable data is completely summarized by two principal components. This plot is related to the most common way of plotting principal components. The scores (the same as our usual principal component scores with variance equal to the eigenvalues) are plotted for specimens as points, and the rows of the column eigenvectors are plotted for each variable as lines or vectors. The properties of this form of the biplot are that the distance between points are exactly Euclidean distances between specimens in original units of measurement, and the two vectors each of length one are at right angles, and have angles with the principal component axes which are the angle whose cosine is the corresponding eigenvector coefficient, that is the amount of rotation of the bivariate ellipse axes best fitting the data to form the new PC axes. The length of each vector is scaled to length 1. Each specimen's value, in terms of deviations from the mean, for each of the two variables can be reconstructed accurately (depending on the accuracy of the plot) by dropping a perpendicular from each specimen point onto each variable vector. This is illustrated for the 6th point in the deviation matrix with CIL = 0.775 and LD = 0.310.

Fig. 1

Biplots of first ten specimens of Zygodontomys from Divedive, using only characters CIL (condylo-incisive length) and LD (length of the diastema).

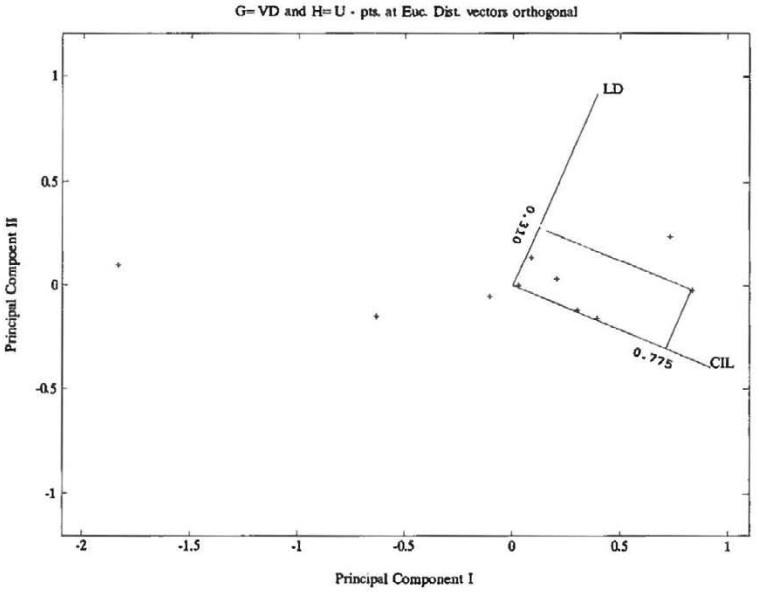


Fig. 1a

Specimens euclidean distance apart, variables as orthogonal vectors.

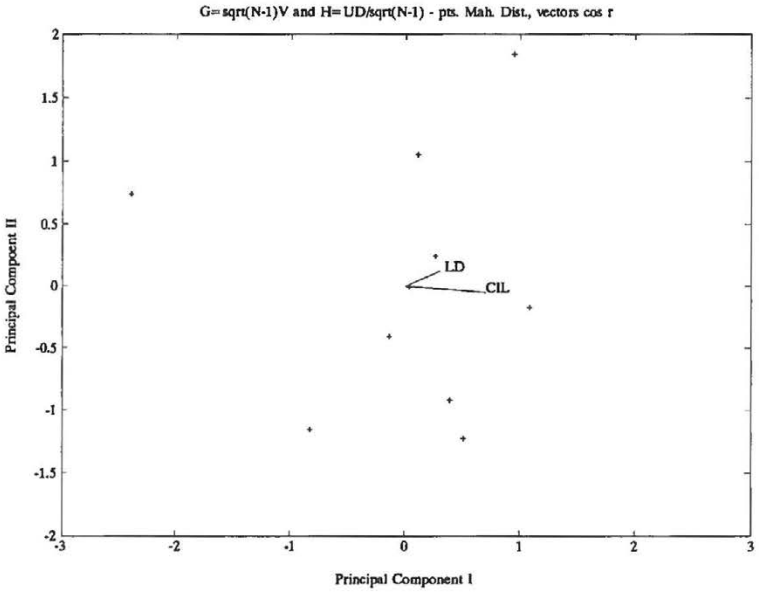


Fig. 1b

Specimens Mahalanobis distance apart, variables at angle arc-cosine of the correlation coefficient.

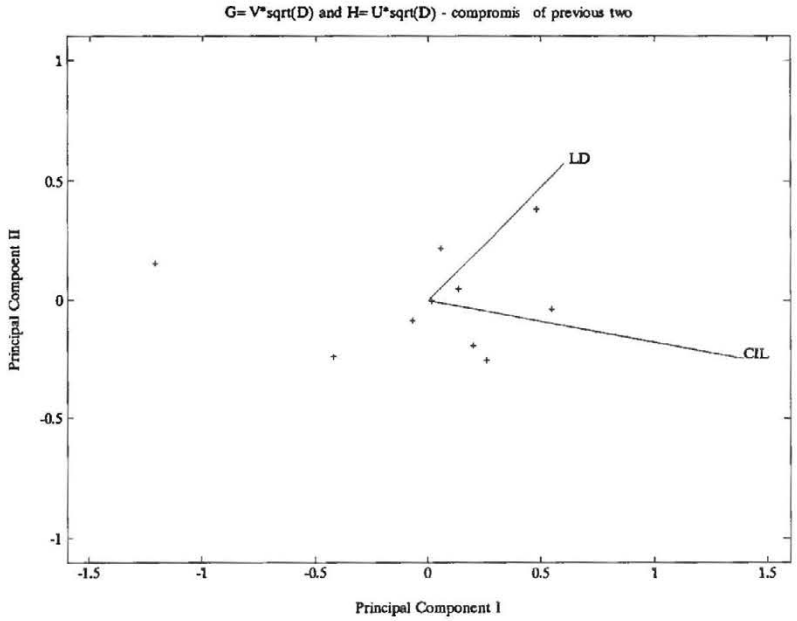


Fig. 1c
Compromise of 1a. and 1b.

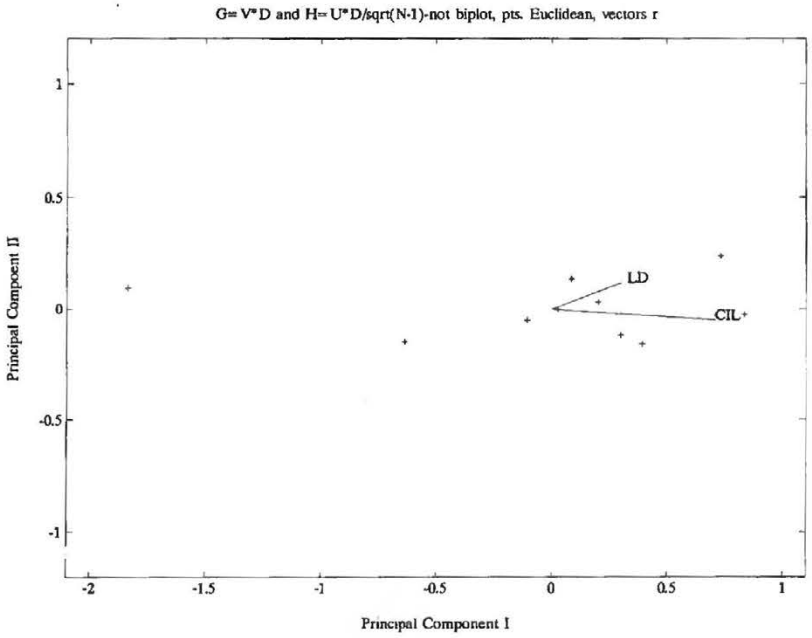


Fig. 1d
Rohlf combination of specimens euclidean distance, variables at arc-cosine correlation - not a true biplot.

The second form of the biplot, figure 1b, also has the property that the projection vector to the specimen point on the variable vector exactly reconstructs the data item for that variable. Now, however, the scores are scaled so that the distance between specimen points is Mahalanobis distance, which involves the inverse of the variance-covariance matrix. Variables with large variances are down weighted in the plot of the specimen points. However, the vectors for each variable have lengths equal to their respective standard deviations, and the angle between the two variables is the angle whose cosine is the correlation coefficient between the two variables.

A compromise biplot is given in figure 3c, but it does not have such simple interpretations in terms of distances and angles. It is a true biplot in that the data is completely recoverable from it. This is the form commonly used in correspondence analysis, where scale is not so important. It is also useful for other applications of the singular value decomposition, such as partial least squares (Bookstein, 1991).

A fourth form which is not a true biplot is favored by Rohlf (pers. comm.). It has the Euclidean distance property of the 1st plot, and the correlation coefficient and standard deviation properties for the variables in the second plot. It is not a true biplot as the projections of the data points onto the variable vectors will not reproduce the data. A most useful additional property of this display is that the distances and lengths of vectors are in original measurement units. Note that this was true only for the observation points in the 1st plot, for the variable vectors in the 2nd plot, and for neither in the third plot.

We will now describe in matrix algebra the verbal descriptions I have given up to now, and then give a number of data examples from my own research.

For any data array Y (usually but not necessarily mean centered), or alternatively Z (which is column standardized by dividing columns by the standard deviation for the respective column, i.e. when one is thinking in terms of correlations), we may decompose Y (or alternatively Z) as:

$$Y=VDU'$$

V is the matrix of scores (actually scaled principal component scores) which have variance 1/(n-1), so V is orthonormal, e.g. V'V=I the identity matrix. If Y has n rows and p columns and is of full rank, then V has n rows and p columns. D is a diagonal matrix which contains the singular values. These are the square roots of the eigenvalues times the square root of n-1, e.g. $D= ((n-1)L)^{1/2}$. Another way of saying the same thing is that they are the standard deviations of the more traditionally scaled principal component scores times the square root of n-1. U is the matrix of column eigenvectors scaled to length 1, which are also orthonormal (U'U=I). Repeating, if Y is of full rank and n is greater than p, then V is n x p, D is p x p and U is p x p. A more compact form is available if

the rank of Y is r<p and then sizes of the matrices are respectively: V is n x r, D is r x r and U is p x r. The usual principal component scores scaled to have variance equal to the eigenvalues are given by:

$$PCScore=VD$$

Incidentally the singular value decomposition of a symmetric matrix such as a covariance matrix or correlation matrix gives D as the matrix of eigenvalues, and V=U the matrix of eigenvectors.

We may also write Y=VDU' in the following informative way:

$$Y=v_1d_1u_1' + v_2d_2u_2' + ...v_id_iu_i'... + v_pd_pu_p'$$

where v_i is the ith column of V, d_i=the ith diagonal element of D, and u_i is the ith column eigenvector in U. We may then, as in principal components summarize and plot our data in terms of the first r principal components. Then the proportion of original variance is:

$$\frac{\sum_{i=1}^{r} d_i^2}{\sum_{i=1}^{p} d_i^2}$$

Using this formulation to develop the biplot we may write (using the notation of Gabriel, 1968):

$$Y=GH'$$

If G contains, as in the first form of the biplot, the first two columns of VD and H the first two columns of U, then we plot the n rows of G for specimens and p rows of H for variables. This plot will be the best two dimensional summary of Y (the mean centered data matrix) in a principal components or least square sense. That part of the data described by the plot can be reconstructed exactly from G and H. G will be the same as the commonly plotted principal component 1 and 2 scores scaled to have variance equal to the eigenvalues. Then as was stated above, the data points G will be separated by that part of the Euclidean distance between the specimens represented by the first two principal components.

The plot of the first two columns of the H matrix will be a plot of the eigenvector coefficients for each variable. If we project the data points onto axes along these vectors, we then see how much each specimen's location in the principal components plot is determined by each variable. The angle of the variable vector

and the display axes will be the cosine of the angle each variable makes with the new axes, or in other words the amount of rotation of the original variable axes to the principal component scores axes. Each data value for each variable will be reconstructed by multiplying the specimen vector in G by the variable vector in H.

An alternative plot in figure 1b lets G=V and H=UD. In this case the points will have distances apart represented by their Mahalanobis distance (actually divided by $(n-1)^{0.5}$), while the variable vectors will have lengths proportional to their standard deviations, and be at angles to each other given by the arc cosine of the correlation coefficient between them. Of course the data values are recovered by the product of the rows of G and H.

Krzanowski (1988) offers a compromise form of biplot, which lets $G=VD^{.5}$ and $H=UD^{.5}$ which balances the contribution of the singular values between the two components. The geometric interpretation is less intuitive in this case. This is the form usually used in correspondence analysis and partial least squares.

Jackson (1991) has generalized the biplot partition by defining $G=VD^{c/2}$ and $H=UD^{(1-c/2)}$. Then for Figure 1a c=0, Figure 1b c=1, and Krzanowski's suggestion shown in Figure 1c has c=.5. Rohlf's compromise doesn't fit this formulation.

One further generalization in Greenacre (1984) includes both the biplot and other similar displays with biplot like properties, such as Rohlf's suggested formulation. Greenacre defines $G=VD^a$ and $H=UD^b$, where a and b are constants. Then a=1 and b=0 gives the first form, a=0 and b=1 the second, a=1/2 and b=1/2 the third - up to now all biplots have a+b=1 and this is similar to Jackson's suggestion. If we let a=1 and b=1, then we no longer have a biplot but rather the approach Rohlf favors with properties discussed earlier.

Biplots for five variables (the deviation matrix given above) for *Zygondontomys* from Dividive is given in figure 2. The V, D, and U matrices are given below for the biplot together with the summary of how much of the trace of the variance is in the biplot (the sum of the first two eigenvalues). Again note that the sum of all five eigenvalues 0.6874 is the same as the sum of variances in the original covariance matrix.

Eigenvalues

Eig.Val.	cum.%	
0.6329	92.0821	
0.0309	96.5741	– note that biplot will summarize 96.6%
0.0146	98.6942	of the total variance
0.0081	99.8694	
0.0009	100.0000	

V =

–0.7935	0.2294
–0.2869	–0.4762
0.1049	0.4168
–0.0267	0.0757
0.0129	0.1397
0.3476	–0.1203
0.3324	0.4139
0.0155	0.0382
0.1192	–0.5550
0.1746	–0.1622

D =

2.3866	0
0	0.5271

U =

CIL	0.8857	–0.3387
LD	0.3838	0.5071
LM	–0.0112	–0.6896
BM1	–0.0164	–0.1360
LIF	0.2603	0.3662

Residual variance covariance matrices and residual data matrices for the *Zygodontomys* data are given below. First the residual covariance matrix:

	CIL	LD	LM	BMI	LIF
CIL	0.0007	–0.0010	–0.0013	–0.0011	–0.0010
LD	–0.0010	0.0059	0.0016	0.0021	–0.0052
LM	–0.0013	0.0016	0.0027	0.0020	0.0024
BM1	–0.0011	0.0021	0.0020	0.0030	0.0010
LIF	–0.0010	–0.0052	0.0024	0.0010	0.0112

Residuals expressed in terms of the correlation matrix are:

	CIL	LD	LM	BM1	LIF
CIL	0.0013	0.0042	0.0142	0.0260	0.0057
LD	0.0042	0.0554	0.0378	0.1048	0.0664
LM	0.0142	0.0378	0.1552	0.2516	0.0743
BM1	0.0260	0.1048	0.2516	0.8039	0.0683
LIF	0.0057	0.0664	0.0743	0.0683	0.1924

Note that quite a bit of the correlation (0.2516) between LM (length of the molar tooth row) and BM1 (breadth of the 1st molar) is not recovered in this two dimensional representation; and only (100(1 - 0.8039)) = 20% of the variance of BM1.

Fig. 2

Biplots of first ten specimens of Zygodontomys from Divedive for three additional characters - LM (length maxillary tooth row), BM1 (breadth of first molar), and LIF (length incisive foramen). a-d. as in figure 1

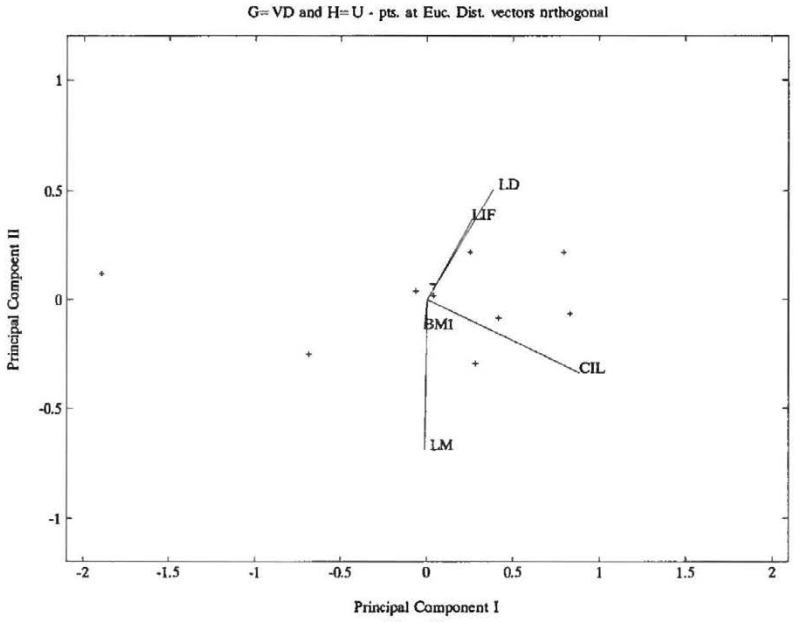


Fig. 2a

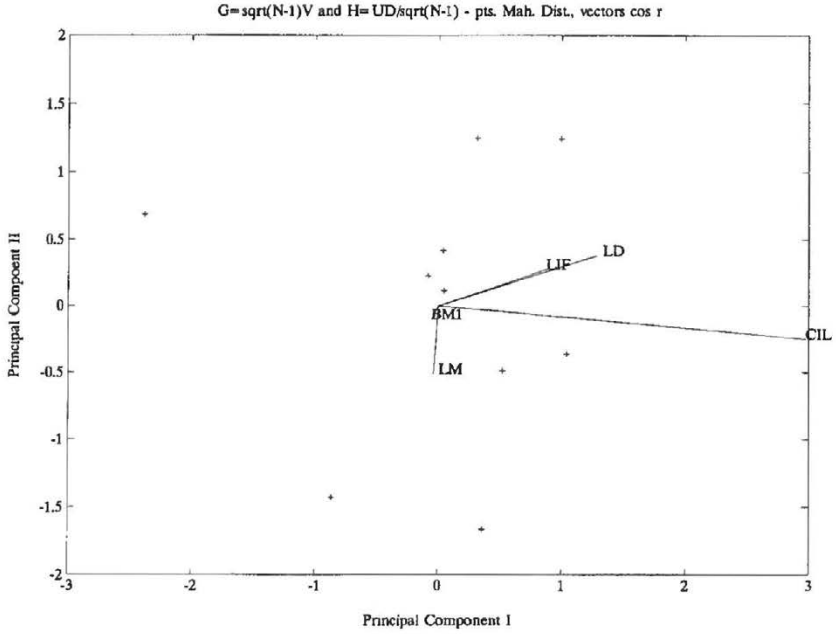


Fig. 2b

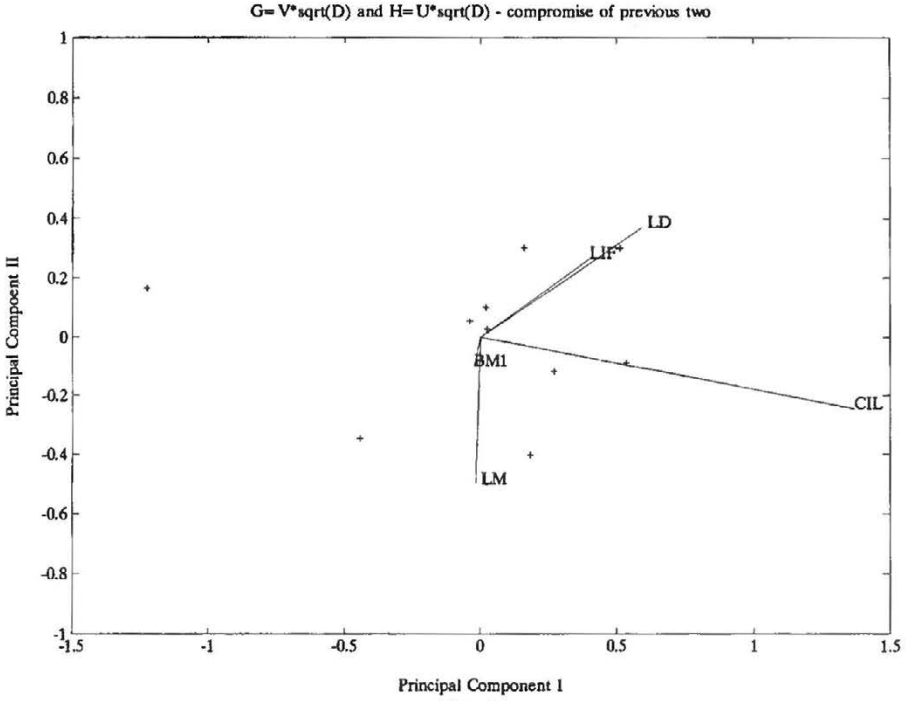


Fig. 2c

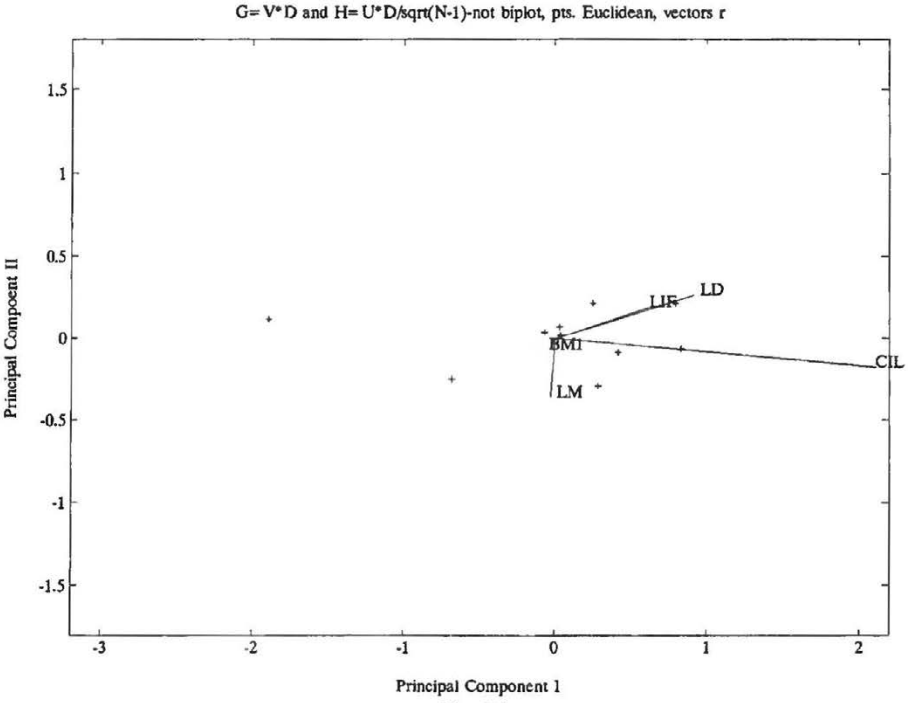


Fig. 2d

Estimated data (u1*d1*v1'+u2*d2*v2' and adding back the means)
Yest =

CIL	LD	LM	BM1	LIF
23.006	6.174	4.042	1.254	5.061
24.203	6.449	4.285	1.285	5.239
24.872	7.047	3.950	1.206	5.655
24.655	6.835	4.078	1.235	5.508
24.727	6.889	4.053	1.229	5.545
25.481	7.126	4.139	1.235	5.702
25.353	7.255	3.945	1.197	5.796
24.751	6.864	4.090	1.236	5.527
25.076	6.800	4.303	1.275	5.476
25.123	6.956	4.159	1.244	5.587

and residual data matrix (subtracting above from original data).

CIL	LD	LM	BM1	LIF
–0.006	0.025	0.007	0.045	–0.011
–0.003	0.000	0.014	–0.035	0.010
0.027	–0.097	–0.050	–0.056	0.044
–0.005	–0.085	0.021	–0.035	0.142
0.022	0.110	–0.053	–0.029	–0.245
0.018	0.023	–0.039	–0.035	–0.102
–0.053	0.094	0.104	0.102	0.053
–0.001	–0.014	0.009	–0.036	0.023
–0.026	0.049	0.046	0.074	0.023
0.027	–0.106	–0.059	0.005	0.062

For CIL all but one residual is less than recording accuracy (0.05 mm.), however for the other variables important parts of the data are not recovered. Here CIL, that character with the most variance, contributes most to the total variance so that it is perhaps misleading to report the results in terms of the raw data. Logarithms were used in the published study (Voss et al. 1990), and then this makes the variance more homogeneous and the residuals better behaved.

The biplots for the complete sample of 12 variables for 68 specimens from Dividive is shown in Figure 3.

The data for Mediterranean birds displayed in figure 3 in Marcus (1990) is here also shown as a biplot in the various formulations (Figure 4 a-d). In this and the complete previous example the first two principal components summarize 85% of the variability, so that residual analysis is appropriate. The distance between points are then not so well described. Parts of the standard deviations are not expressed, and the angles between variables are projections into this two dimensional plot. For the *Zygondontomys* example, as analyzed in Voss et al. (1990) where the logs of 12 variables are used, it is difficult to interpret principal components past the first two. They seem to represent measurement error and unpatterned residual variation, so that the biplot may be a fair display of the important biological variation in the sample for the logarithms of the data.

Fig. 3

Biplots of all 68 specimens of Zygodontomys from Dividive for all 12 cranial characters. (Voss et al. 1990). a-d. as in figure 1. Additional characters BR (breadth of rostrum), BPB (breadth of palatal bridge), BZP (breadth of zygomatic plate), BB (breadth of braincase), LIB (least orbital breadth), DI (depth of incisor), and LOF (length of orbital fossa)

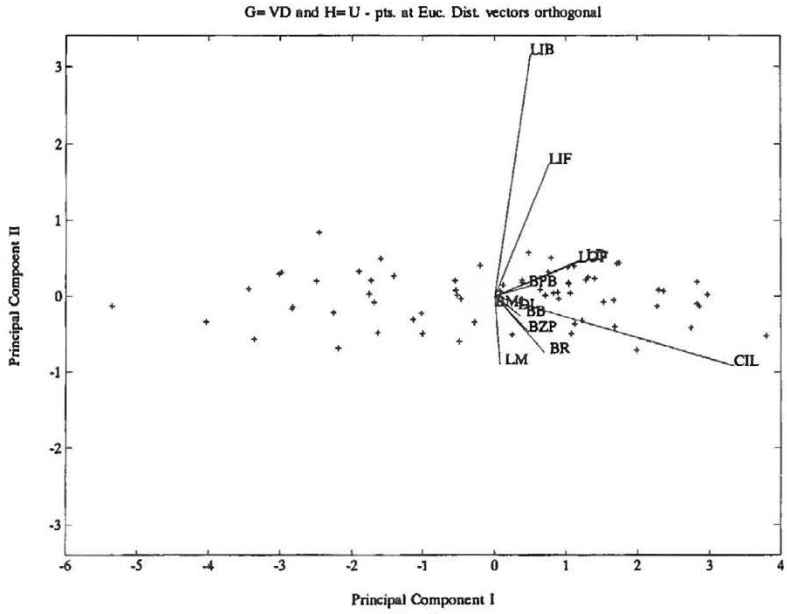


Fig. 3a

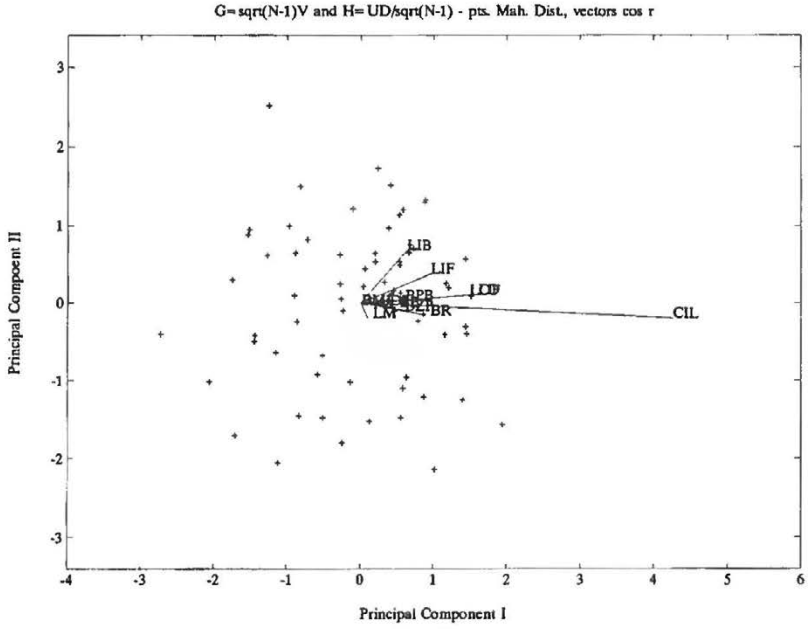


Fig. 3b

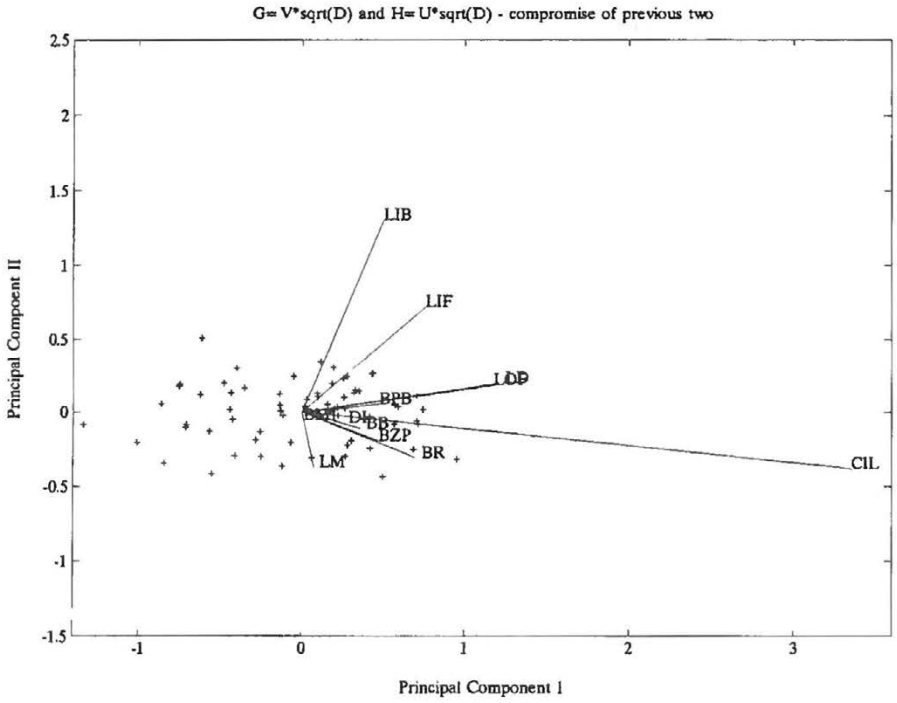


Fig. 3c

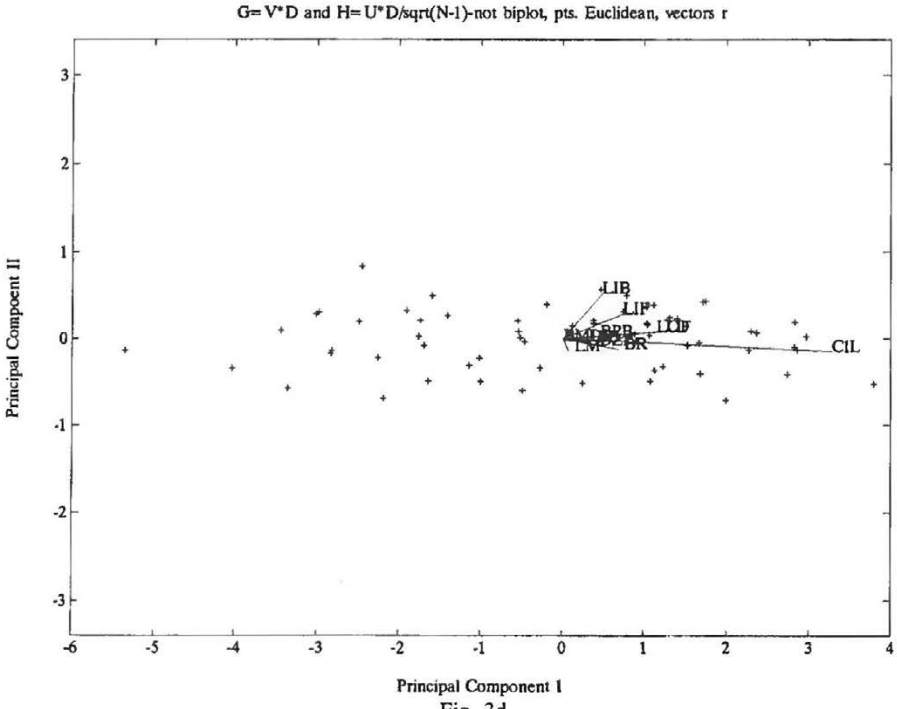


Fig. 3d

Fig. 4

Biplots of Mediterranean birds (Blondel et al., 1984) for 126 species and 7 morphological distances and weight. a-d. as in figure 1. Characters logs of lengths of Wing, Tail, Culmen,Tarsus, and Midtoe; Billh (height), Billw (width); and Weight

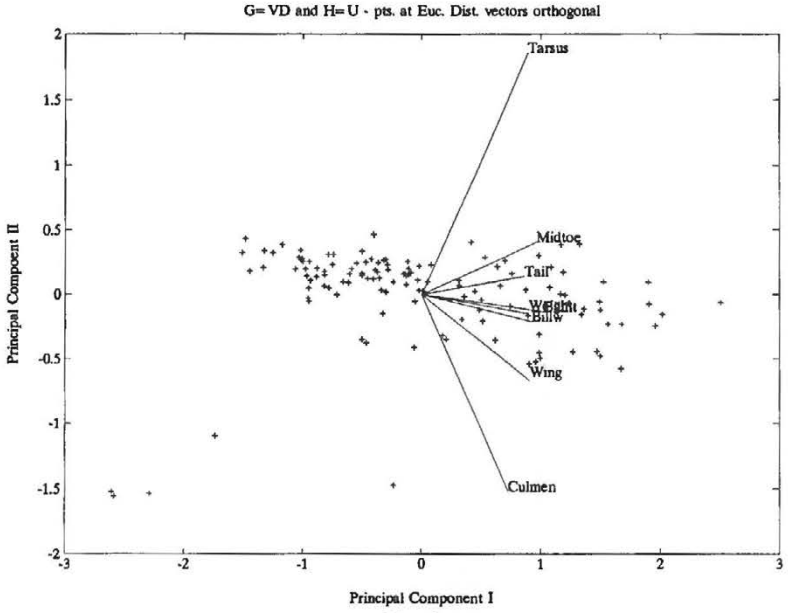


Fig. 4a

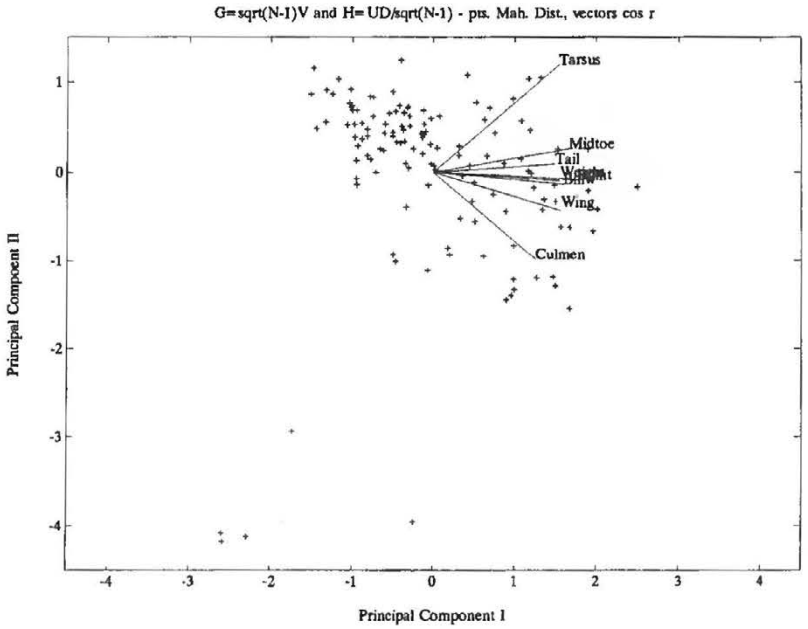


Fig. 4b

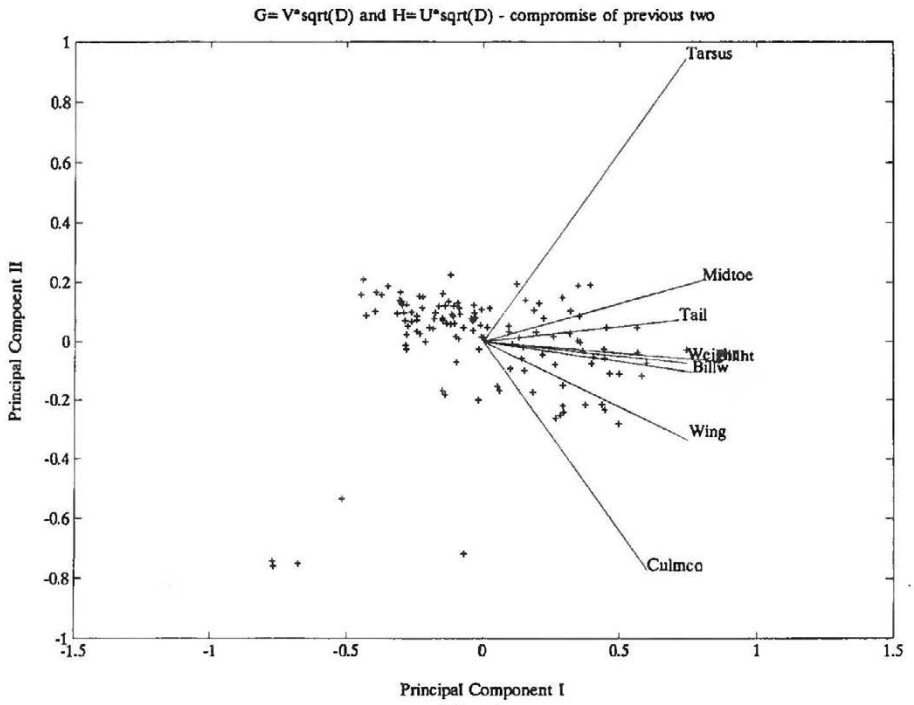


Fig. 4c

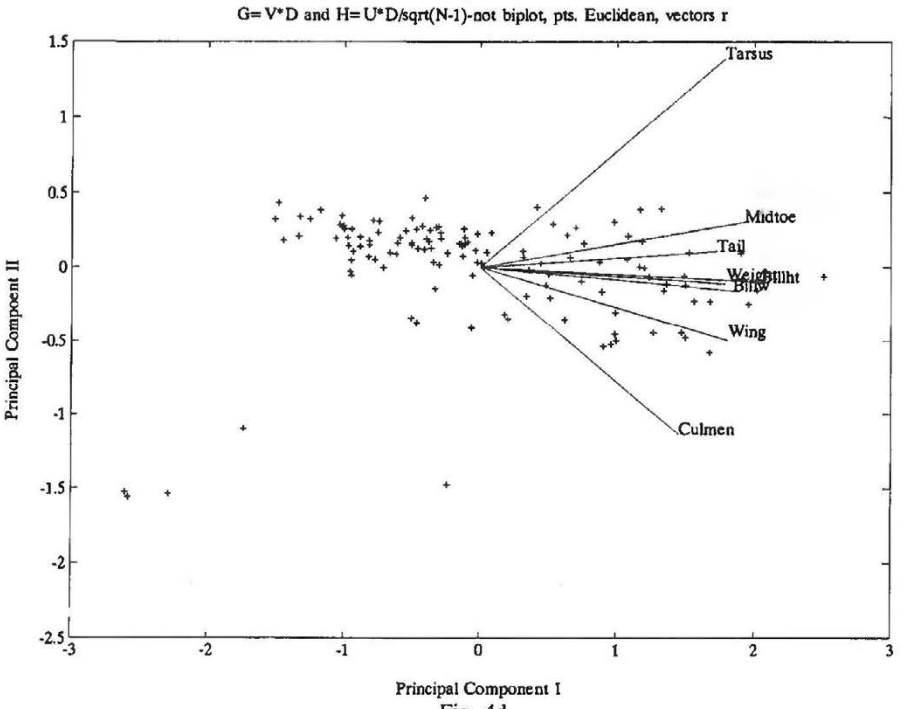


Fig. 4d

For the Mediterranean birds, in figure 4a the angles are the angles of the variable vectors to the respective component axes, but in 2b the angles between vectors only approximate the correlations between variables. Also as was pointed out in Blondel et al. (1984) and Marcus (1990) certain values for certain birds are not well expressed by the two dimensional principal components plots, and thus for the biplots. An analysis of residuals was necessary to reveal the values not well summarized by the bivariate plots.

The five species in the lower left of the plot are the hummingbirds and figure 4 shows how this plots shows that their short legs (tarsus length) and long bills (culmen length) separates them from the other birds in the data set. Further details are discussed in Blondel et al. (1984).

We can see that the compromise solution of figures 3c and 4c are intermediate between 3a and 3b, and 4a and 4b respectively.

Recall that the Rohlf alternative to the biplot display has useful properties not present in the first two formulations, but is not a true biplot in that the data cannot be approximated directly from the display. To repeat, Rohlf suggests plotting the row vectors of G=VD and H=UD (note that in this case Y is not equal to VDU'). I have modified this formulation further as $G=VD*(n-1)^{0.5}$ and $H=f*UD*(n-1)^{-0.5}$ so that the points then have the property of being a distance apart equal to their Euclidean distance, and the vectors for variables will have lengths equal to their standard deviations. The fudge factor, f=1, adjusts the vector lengths and may produce a more pleasing plot, but destroys the interpretation of the length of the vector and the exact graphical reproduction of the data. This can be compensated for by dividing through by f if known from the plot. The f multiplier does not affect the angles between variable vectors which are equal to the arc cosine of the correlation coefficients. Another feature of Rohlf's proposal is that if the data is measured in the same units for all variables (for example millimeters), then the biplot axes will be in those units and the points can be measured from each other in mms. The vectors corresponding to the variables will also have lengths equal to f standard deviations in millimeters. The data can be reconstructed from the specimen and point vectors by dividing their inner product by the length of the variable vector. When one has mixed scales, and standardizes the data to be scale free this may be a less important consideration. Then the symmetrical Krzanowski choice may be appropriate as pointed out above.

Another example is given, which also makes some useful points about the use of principal components in systematic studies. Stuenes and Marcus (1991) in a comparison of four species of hippopotami, two living and two recently extinct, used data on skull dimensions transformed to logarithms. The names of the living species with their plotting symbols are *Hippopotamus amphibius* (a) and *Hexaprotodon liberiensis* (b), and the subfossil species are *Hippopotamus lemerlei* (l) and *Hippopotamus madagascariensis* (m). Ten measurements were taken (see Figure 5 legend for a list). Only the Rohlf biplot like graphs are shown

Fig. 5

Biplot of 4 species of hippopotamus for cranial variables. a. Rohlf form of biplot for best two diagnostic characters as logs of lengths - LO (orbit); LEI (face from the orbit to eye socket). b. Rohlf form of biplot for all 10 cranial variables. Additional logged characters - LPO (length of skull), LPM (length premolar molar tooth row), LOO (width at auditory meatus), LEE (width across postorbital processes), WZZ (maximum bizygomatic width), LWP (largest palatal width), WAC (muzzle width at canines), and LWR (least width of rostrum). Note two fossil species are separated better on a. then b. Plotting symbols are a= Hippopotamus amphibius, b = Hexaprotodon liberiensis, l = Hippopotamus lemerlei, m = Hippopotamus madagascariensis

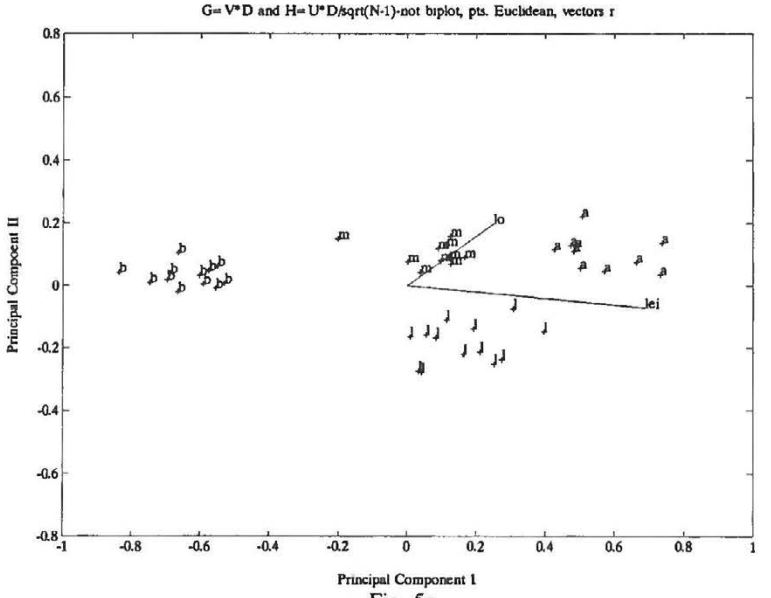


Fig. 5a

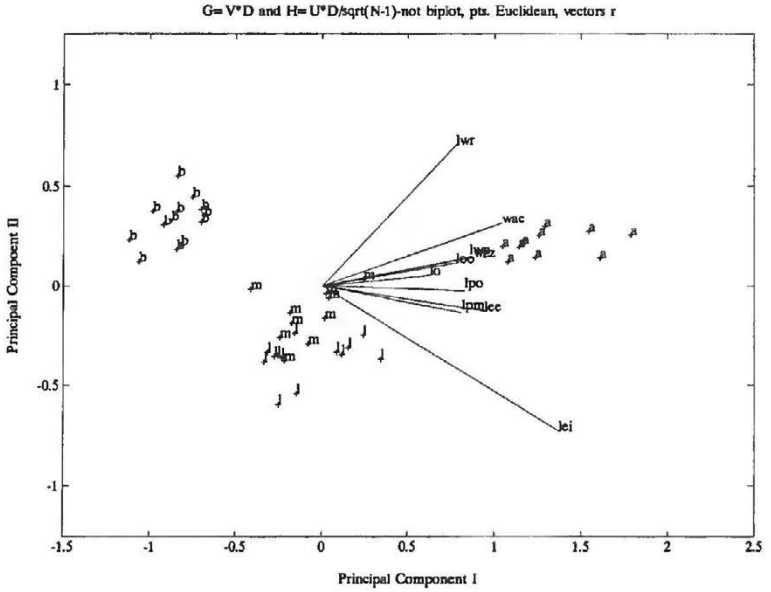


Fig. 5b

here for two choices of variables in figure 5. First an exact biplot of the two best discriminating variables for the 4 species is given, and then the biplot for all 10 variables included in the study.

Note that the 4 species form separate clusters on a biplot based on length of the orbit (lo) and length of the face (lei), while the biplot for all 10 variables, shows less distinction between the two Madagascan sub-fossil species (Figs. 5a and 5b). This illustrates an important point. An overall principal component plot of all of the variables measured may not ordinate the taxa distinctly in the plane of the first two principal components even when an overwhelming, in this case 95.6%, percentage of the variability is represented by these components. In this case the logs of the data were used because of the difference in variance of the lengths used, and the large difference in size of the various species. Cross-validation (Krzanowski, 1987; also see Reyment, 1991) did a good job of finding the structure for the two diagnostic variables, but also suggested a third one. Adding this third variable gives two dimensional display with more overlap of the extinct species.

Some relevant data for the Hippo example (logged data) are:

	lpo	lo	lei	lpm	loo	lee	wzz	lwp	wac	lwr
Means	6.0931	4.4897	3.9355	5.1424	5.1051	5.2570	5.5745	4.6987	5.2151	4.1741
s	0.2322	0.4323	0.1996	0.2331	0.2213	0.2689	0.2496	0.2454	0.3098	0.2962

Eigenvalues

Values	Cum.%
8.6800	86.7995
0.8252	95.0512
0.2539	97.5907
0.0829	98.4195
0.0503	98.9221
0.0372	99.2942
0.0309	99.6034
0.0171	99.7742
0.0119	99.8930
0.0107	100.0000

– Note that more that 95% of the variance is summarized by the biplot.

r =

	lpo	lo	lei	lpm	loo	lee	wzz	lwp	wac	lwr
lpo	1.0000	0.8874	0.8174	0.9586	0.9561	0.9540	0.9532	0.9446	0.9212	0.7062
lo	0.8874	1.0000	0.7155	0.9187	0.7846	0.9067	0.7782	0.7730	0.6896	0.3461
lei	0.8174	0.7155	1.0000	0.8005	0.8580	0.8448	0.8916	0.8664	0.8152	0.6465
lpm	0.9586	0.9187	0.8005	1.0000	0.9191	0.9589	0.9123	0.9161	0.8477	0.5984
loo	0.9561	0.7846	0.8580	0.9191	1.0000	0.9371	0.9830	0.9685	0.9523	0.7940
lee	0.9540	0.9067	0.8448	0.9589	0.9371	1.0000	0.9330	0.8998	0.8780	0.6106
wzz	0.9532	0.7782	0.8916	0.9123	0.9830	0.9330	1.0000	0.9692	0.9519	0.8059
lwp	0.9446	0.7730	0.8664	0.9161	0.9685	0.8998	0.9692	1.0000	0.9260	0.8100
wac	0.9212	0.6896	0.8152	0.8477	0.9523	0.8780	0.9519	0.9260	1.0000	0.8530
lwr	0.7062	0.3461	0.6465	0.5984	0.7940	0.6106	0.8059	0.8100	0.8530	1.0000

residual r =

	lpo	lo	lei	lpm	loo	lee	wzz	lwp	wac	lwr
lpo	0.0307	0.0130	0.0519	0.0035	0.0023	0.0081	0.0059	0.0027	0.0108	0.0129
lo	0.0130	0.0240	0.0271	0.0041	0.0068	0.0094	0.0064	0.0013	0.0011	0.0191
lei	0.0519	0.0271	0.2112	0.0425	0.0177	0.0072	0.0136	0.0014	0.0307	0.0415
lpm	0.0035	0.0041	0.0425	0.0388	0.0014	0.0049	0.0065	0.0097	0.0061	0.0126
loo	0.0023	0.0068	0.0177	0.0014	0.0239	0.0047	0.0033	0.0004	0.0003	0.0081
lee	0.0081	0.0094	0.0072	0.0049	0.0047	0.0327	0.0018	0.0191	0.0086	0.0008
wzz	0.0059	0.0064	0.0136	0.0065	0.0033	0.0018	0.0163	0.0037	0.0060	0.0085
lwp	0.0027	0.0013	0.0014	0.0097	0.0004	0.0191	0.0037	0.0377	0.0224	0.0005
wac	0.0108	0.0011	0.0307	0.0061	0.0003	0.0086	0.0060	0.0224	0.0508	0.0097
lwr	0.0129	0.0191	0.0415	0.0126	0.0081	0.0008	0.0085	0.0005	0.0097	0.0287

HOTELLING'S T^2 AND STUDENT'S t

The test statistic used to test the null hypothesis that two populations mean vectors, or centroids are the same, is Hotelling's T^2, the multivariate extension of the ordinary two sample t. The assumptions of normality, and equality of variance must be extended to multivariate normality and homogeneity or equality of variances to homogeneity of variance-covariance matrices.

The formula for two sample t for n_1 individuals in sample 1 and n_2 in sample 2 for the univariate case is:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{x}_1 - \bar{x}_2}}$$

where t has student's t distribution when the hypothesis of 0 difference between the means is true, with n_1+n_2-2 degrees of freedom. $\bar{X}_1$ and $\bar{X}_2$ are the means for the two samples, $S_{\bar{x}_1 - \bar{x}_2}$ is the standard error of the mean difference, and s^2 is the pooled within sample variances for the two samples based on n_1+n_2-2 degrees of freedom.

t squared is written:

$$t^2 = \frac{(\bar{X}_1 - \bar{X}_2)^2}{S_{\bar{x}_1 - \bar{x}_2}^2} = \frac{n_1 n_2}{n_1 + n_2} \frac{(\bar{X}_1 - \bar{X}_2)^2}{S^2}$$

and has an F distribution with 1 degree of freedom for the numerator and n_1+n_2-2 for the denominator.

We may rewrite the last equation as:

$$t^2 = (\bar{X}_1 - \bar{X}_2) \frac{n_1 n_2}{n_1 + n_2} (S^2)^{-1} (\bar{X}_1 - \bar{X}_2)$$

Then the formula for Hotelling's T^2 will look very similar. It may be written in an analogous way as:

$$T^2 = (\bar{X}_1 - \bar{X}_2)' \frac{n_1 n_2}{n_1 + n_2} S^{-1} (\bar{X}_1 - \bar{X}_2)$$

Where $\bar{X}_1$ and $\bar{X}_2$ are now vectors of means, ' the transpose operator, and S is the pooled within variance covariance matrix of the variables, and -1 the inverse operator, is analogous to a univariate reciprocal.

Mahalanobis distance squared is a closely related statistic and is usually written as:

$$D^2 = (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2)$$

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} D^2$$

Therefore Hotelling T^2 written in terms of Mahalanobis D^2 is:

If g samples are used to estimate the pooled within covariance matrix, with $N = n_1 + n_2 + ... + n_g$ total observations then T^2 is still the same with i and j replacing 1 and 2 as the subscripts for any specific comparison of the ith and jth centroid. Therefore the D^2 values in the output of most multivariate packages may be converted to T^2 values by this last formula.

We use an F statistic to test our hypothesis of equality of centroids with two samples and the formula for F is:

$$F = \frac{n_1 + n_2 - p - 1}{(n_1 + n_2 - 2)\, p} T^2$$

[see Morrison and other references to textbooks mentioned earlier] which has an F distribution with p, and n_1+n_2-p-1 degrees of freedom, for testing the null hypothesis of equal mean vectors.

If the variance-covariance matrix is estimated from pooled samples across g samples then,

$$F = \frac{N - g - p + 1}{(N - g)\,p}\ T^2$$

for N total observations in g groups. This F then has p and N–g–p+1 degrees of freedom.

Note that these formulae reduce to the usual $F=t^2$ for p=1 as then n_1+n_2-2 cancels out. We may also find a confidence interval for the difference between two centroids, though this has not been used much in morphometrics (Morrison, 1991). The confidence interval for the mean difference does not give a confidence interval for the distance squared Δ^2 that Mahalanobis D^2 estimates as that confidence interval also depends on the unknown covariance matrix. A program for finding approximate confidence intervals for the parameter Mahalanobis Δ^2 is given in the Michigan Morphometrics workshop disks; where it is also pointed out that D^2 is a biased estimate of Δ^2.

That bias correction can become quite important for small samples as the unbiased estimate may become negative. In this case the estimate is set to 0. An example is given for some mole data samples (discussed in this volume by Loy, Corti and Marcus) where the sample sizes vary a great deal. Very misleading D^2 values arise for small samples, and large numbers of variables that are frequently encountered in morphometrics. The formula for the unbiased estimate is:

$$Du^2 = \frac{n_1 + n_2 - p - 3}{n_1 + n_2 - 2}\ D^2 - \frac{(n_1 + n_2)\,p}{n_1\,n_2}$$

for two samples. Note that there must be more than p+3 individuals in the two samples combined to get a non-negative unbiased estimate of Δ^2.

For several samples where the degrees of freedom are pooled over samples, for a total sample size of $N=n_1 + n_2 + ... + n_g$, and g groups or samples, then for each distance comparison from sample i to j.

$$Du_{ij}^2 = \frac{N - g - p - 1}{N - g}\ D^2 - \frac{(n_i + n_j)\,p}{n_i\,n_j}$$

The mole example (see Loy *et al.* this volume) gives extreme examples of the affect of bias in Mahalanobis D^2. Larger samples would have reduced the

bias, but they were not available. This matrix was computed using Bookstein coordinates for 13 landmarks, using landmark 10 and 13 as the base, so that there were 11 pairs of x and y coordinates, or 22 variables. However we see that for samples of 1 and 4, comparing male and female min specimens the bias reduces D^2 from 41.4 to 8.73. All negative values have been set to 0. For the larger sample comparisons, for example between the teub females and the teus females the bias correction reduces D^2 from 30.7 to 24.4.

ANALYSIS OF VARIANCE AND MULTIVARIATE ANALYSIS OF VARIANCE

Tests of hypotheses in analysis of variance are based on the ratio of two Mean Squares. For example the test for locality differences for the *Zygodontomys* data for any one variable is a one way analysis of variance.

F in that case is:

$$F = \frac{MS_A}{MS_W}$$

where MS_A is the among sample mean square and MS_W is the pooled mean square within samples or as

$$F = \frac{(N - g)\, SS_A}{(g - 1)\, SS_W}$$

where SS stands for the sum of squares about the mean.

We may rewrite F as:

$$F = \frac{N - g}{g - 1}\; SS_W^{-1}\; SS_A$$

We reject the null hypothesis of equal locality means if F is larger than F with our chosen significance level, and g-1 and N-g degrees of freedom.

If we let SS_A be the multivariate among sums of squares and cross products matrix with the univariate sums of squares on the diagonal, and SS_W the sum of squares and cross-products pooled over groups (localities in this case) then

canonical variates, multiple discriminant analysis and Multivariate Analysis of Variance are all functions of

$$SS_W^{-1}SS_A$$

and we see the formulation of the statistic is the same as in univariate analysis except for the factor (N-g)/(g-1) because we are using sums of squares rather than variance-covariance matrices. Mahalanobis D^2 between groups is a function of SS_W as well and may be written for group i and j as:

$$D^2 = (\bar{X}_1 - \bar{X}_2)' \ (N - g) \ SS_W^{-1} \ (\bar{X}_1 - \bar{X}_2)$$

However the comparison is not so simple as for t^2 and T^2.

$SS_W{}^{-1}SS_A$ is no longer a scaler and we have a matrix of results, which must be used to test our multivariate hypothesis about centroids. All of the proposed test statistics are based on the eigenvalues of $SS_W{}^{-1}SS_A$, and not surprisingly some are functions of all of the Mahalanobis D^2. One of them, is based on the largest eigenvalue. The likelihood ratio test is yet another. Two other statistics are based on the sum of the eigenvalues. A comparison of the tests should depend on the power of the test, but this is very difficult to compute and no one test is uniformly most powerful, but each is better against certain kinds of alternative hypotheses. Seber (1984) suggests that for small departures from the null hypothesis the order of the tests in terms of decreasing power is Pillai's trace, the likelihood ratio test, the Lawley-Hotelling Trace, and finally the test based on the largest root. However, if most of the differences are concentrated in one dimension (in terms of canonical variates) then the powers are in the reverse order. Note also that this matrix product is frequently not of full rank and the number of eigenvalues will be equal to the lesser of the number of variables (p) and groups minus one (g-1).

For the mole example, testing for sex by locality interaction in a two way analysis of variance for several localities, the largest root test was highly significant while the other three tests gave very large p values. There was no other evidence for sexual dimorphism for any of the localities, or in general (Loy et al., this volume).

Multiple comparison procedures can be very complicated in multivariate analysis of variance as there are very many ways in which samples can differ, on one or more variables. Tests are suggested for consecutive roots, but Harris (1975) casts doubts on these.

A priori comparisons or contrasts among centroids for all variables are much more straightforward and involve pre-multiplying the matrix of means by a matrix of contrast vectors, assuming groups in rows, and variables in columns. This is

essentially the same as contrasts in univariate analysis of variance for the differences between means, and the vectors of multipliers are the same (Sokal and Rohlf, 1981). Comparisons or selections of variables requires post multiplication of the matrix of means by a matrix of constants.

Multiple comparison procedures on all centroids are like the problem in univariate analysis of variance. A comparable approach to Bonferroni adjustments to all possible t tests, based on the pooled within mean squared error, is the following. Test all possible mean differences using D^2 or T^2 statistics and corresponding F tests (see above). Adjust the significance level used in the F table, by dividing the nominal significance level α by g(g-1)/2 for an overall α significance level, and only reject mean differences if F exceeds the tabled F value for this α /(g(g-1)/2) probability. For example using an $\alpha = 0.05$ test for 10 localities, enter the F table with 0.05/45 ~ 0.001 and reject the hypothesis for all F values exceeding the 0.001 tabled F value. Note that this procedure is conservative and less powerful than some others, but it is very easy to apply. I frequently see no adjustment in the use of the F tests for tables of D^2 values, and this procedure leads to the same much inflated overall significance level as using all possible t tests in the analysis of variance (Snedecor and Cochran, 1967).

In the mole data none of the sex differences are significant (these are every other entry just below the main diagonals of .000's) within species. Here we have 20 samples, and therefore me must divide 0.05 by 190 to obtain an overall $\alpha = 5$ % significance level using the Bonferroni adjustment. We see that comparisons like that between min females and troc males are not significant as 0.004 is greater than the required 0.05/190= 0.00026. If we were only going to compare the sexes we could use 0.005, dividing the nominal alpha level of 0.05 by 10, the number of locality samples with both males and females.

ACKNOWLEDGEMENTS

The hospitality of Antonio Valdecasas, Elisa Bello and Jose Becerra on my visits to Madrid, where this manuscript was completed was most appreciated. F. James Rohlf and Fred Bookstein gave helpful criticisms and suggestions, which I believe improved the paper.

Table 1

Species and sexual discrimination based on Mahalanobis D^2

First Row is Species Code, Second row is Sex and Third row is sample size. Mahalanobis D^2 above the diagonal and unbiased D^2 below the diagonal

min	min	pbr	pbr	tca	tca	teub	teub	teun	teun	teus	teus	toc	toc	troc	troc	tro	tro	tst	tst
f	m	f	m	f	m	f	m	f	m	f	m	f	m	f	m	f	m	f	m
1	4	6	3	3	3	14	13	4	7	11	6	3	3	1	4	10	14	2	1
.000	41.4	120	149	135	130	64.0	63.0	38.3	37.3	71.3	79.4	67.0	53.8	112	84.1	102	107	143	140
8.73	.000	111	137	114	112	94.8	94.8	50.0	45.2	86.0	98.7	53.2	43.3	108	91.3	85.1	108	138	132
72.0	79.4	.000	12.3	149	148	164	153	135	120	173	172	103	101	157	114	73.2	86.6	119	140
91.7	96.6	.162	.000	151	153	179	161	150	131	176	179	128	126	186	133	96.9	110	156	181
80.6	78.2	108	106	.000	.875	82.1	80.0	89.5	77.3	94.7	93.3	47.4	45.6	62.2	57.7	88.4	102	113	97.6
77.1	77.2	107	107	0	.000	75.8	75.0	85.6	74.4	93.0	89.9	43.9	43.6	56.4	52.7	84.8	97.3	108	94.6
29.9	68.3	124	133	56.7	51.8	.000	3.18	30.9	33.2	35.1	42.0	75.4	59.3	72.7	59.1	92.1	97.1	123	131
29.0	68.2	115	119	55.0	51.0	0	.000	30.7	29.4	30.7	36.3	75.6	61.5	75.7	56.8	89.2	92.6	127	136
6.31	29.7	97.8	107	59.2	56.1	18.2	17.9	.000	8.29	30.9	38.5	55.4	43.0	98.7	73.9	88.7	102	148	142
7.53	28.0	87.9	94.0	51.6	49.4	22.0	18.9	0	.000	28.4	32.1	54.1	41.3	89.3	62.1	78.0	89.0	137	138
35.3	61.1	131	130	66.3	65.0	24.4	20.9	17.8	17.8	.000	7.51	83.7	59.2	106	78.0	112	123	199	182
40.1	69.6	129	131	63.7	61.1	28.5	23.9	22.3	19.3	1.00	.000	78.9	64.6	95.2	69.1	102	109	190	172
27.2	30.7	71.3	88.2	24.5	21.8	51.5	51.6	32.4	33.4	57.6	52.5	.000	18.2	49.8	37.8	46.7	54.2	70.5	51.3
16.9	22.9	69.5	86.1	23.1	21.6	38.8	40.5	22.7	23.3	38.4	41.2	1.64	.000	49.9	38.4	41.6	55.8	78.9	65.5
50.0	61.3	101	121	23.5	18.9	36.7	39.0	53.7	48.4	62.5	52.6	13.7	13.8	.000	14.0	47.3	49.5	72.3	89.7
42.2	62.2	81.7	93.1	34.2	30.3	40.3	38.3	48.5	41.3	54.7	46.3	18.6	19.0	0	.000	25.1	27.7	67.4	73.9
59.4	60.1	52.4	67.8	61.1	58.3	69.0	66.7	62.9	56.6	83.9	75.2	28.5	24.4	16.2	13.0	.000	6.03	45.7	53.9
63.6	79.0	63.4	78.4	72.1	68.7	73.5	69.9	73.6	65.8	93.3	80.9	34.8	36.1	18.5	15.6	1.47	.000	47.1	59.1
83.8	94.3	81.1	107	72.9	69.2	86.1	88.5	102	95.4	145	137	39.5	46.1	28.2	38.7	24.4	26.1	.000	27.8
72.2	80.2	87.5	116	51.3	48.9	82.3	86.0	87.4	86.7	122	113	14.9	26.1	32.4	34.3	21.4	26.0	0	.000

F above diagonal and Probability of exceeding F below

1.00	1.40	4.36	4.75	4.30	4.15	2.53	2.48	1.30	1.38	2.78	2.89	2.13	1.71	2.38	2.85	3.95	4.24	4.05	2.98
.151	1.00	11.3	9.98	8.27	8.18	12.5	12.3	4.24	4.88	10.7	10.1	3.87	3.15	3.68	7.75	10.3	14.3	7.83	4.50
.000	.000	1.00	1.04	12.7	12.6	29.2	26.6	13.7	16.4	28.5	21.9	8.74	8.54	5.72	11.6	11.6	15.4	7.60	5.08
.000	.000	.423	1.00	9.63	9.73	18.8	16.7	10.9	11.7	17.6	15.2	8.18	8.01	5.92	9.66	9.49	11.5	7.95	5.75
.000	.000	.000	.000	1.00	.056	8.61	8.28	6.51	6.89	9.48	7.92	3.02	2.90	1.98	4.20	8.66	10.7	5.76	3.11
.000	.000	.000	.000	1.00	1.00	7.95	7.76	6.23	6.63	9.31	7.63	2.79	2.78	1.79	3.83	8.30	10.2	5.52	3.01
.002	.000	.000	.000	.000	.000	1.00	.908	4.08	6.57	9.17	7.50	7.91	6.21	2.88	7.80	22.8	28.9	9.17	5.18
.003	.000	.000	.000	.000	.000	.574	1.00	3.98	5.67	7.77	6.33	7.82	6.36	2.98	7.37	21.4	26.5	9.32	5.35
.209	.000	.000	.000	.000	.000	.000	.000	1.00	.895	3.85	3.92	4.03	3.13	3.35	6.27	10.8	13.4	8.40	4.81
.161	.000	.000	.000	.000	.000	.000	.000	.590	1.00	5.16	4.40	4.82	3.68	3.32	6.71	13.6	17.6	9.05	5.13
.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.00	1.24	8.37	5.92	4.13	9.71	24.8	32.1	14.3	7.08
.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.252	1.00	6.70	5.48	3.46	7.03	16.3	19.4	12.1	6.26
.011	.000	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000	1.00	1.16	1.58	2.75	4.58	5.68	3.59	1.63
.052	.000	.000	.000	.001	.001	.000	.000	.000	.000	.000	.000	.314	1.00	1.59	2.79	4.07	5.85	4.02	2.09
.004	.000	.000	.000	.020	.039	.001	.000	.000	.000	.000	.000	.083	.082	1.00	.476	1.83	1.96	2.05	1.90
.001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.001	.001	.965	1.00	3.04	3.65	3.82	2.51
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.035	.000	1.00	1.49	3.23	2.08
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.021	.000	.113	1.00	3.50	2.34
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.015	.000	.000	.000	1.00	.785
.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.070	.013	.026	.003	.014	.005	.717	1.00

REFERENCES

BLONDEL, J., F. VUILLEUMIER, L. F. MARCUS & E. TEROUANNE, 1984. Is there ecomorphological convergence among Mediterranean bird communities in Chile, California, and France. In *Evolutionary biology,* 18: 141-213 (M. K. Hecht, B. Wallace, and G. T. Prance, eds.). Plenum Publishing: New York.

BOOKSTEIN, F., 1978. The Measurement of biological shape and shape change. *Lecture Notes in Biomathematics,* v. 24. Springer: Berlin.

-1990. Higher order features of shape. Pp. 237-250. In Rohlf, F. J. & F. Bookstein, eds. *Proceedings of the Michigan Morphometrics Workshop.* Spec. Pub. 2, The University of Michigan Museum of Zoology: Ann Arbor.

-1991. *Morphometric Tools for Landmark Data.* Cambridge University Press: New York.

DILLON, W. R. & M. GOLDSTEIN, 1984. *Multivariate analysis. Methods and applications.* John Wiley & Sons: New York.

GABRIEL, K. R., 1968. The biplot graphical display of matrices with application to principal component analysis. *Biometrika,* 58: 453-467.

GREENACRE, M. J., 1984. *Theory and Applications of Correspondence Analysis.* Academic Press: London.

HARRIS, R. J., 1975. *A Primer of Multivariate Statistics.* Academic Press: New York.

HINTZE, J. L., 1987. *Number Cruncher Statistical System.* Version 5:01. Utah.

JACKSON, J. E., 1991. *A User's Guide to Principal Components.* John Wiley & Sons: New York.

KRZANOWSKI, W. J., 1987. Cross-validation in principal component analysis. *Biometrics,* 43: 575-584.

-1988. *Principles of Multivariate Analysis.* Oxford Science Publications: England.

MARCUS, L. F., 1990. Traditional Morphometrics. Pp. 77-122. In F. J. Rohlf & F. Bookstein, eds., *Proceedings of the Michigan Morphometrics Workshop.* University of Michigan Museums: Ann Arbor.

MARCUS, L. F. & M. CORTI, 1989. *Data Analysis in Systematics.* A manual for workshop W3, Fifth International Theriological Congress, Rome, Italy.

MORRISON, D. F., 1991. *Multivariate Statistical Methods.* 3rd. ed., McGraw Hill Inc.: New York.

REYMENT, R., 1990. Reification of classical multivariate statistical analysis in morphometry. In F. J. Rohlf & F. L. Bookstein (Eds.) *Proceedings of the Michigan Morphometrics Workshop.* Special Publication No. 2, University of Michigan Museum of Zoology: Ann Arbor.

-1991. *Multidimensional Paleobiology.* Pergamon Press: Oxford.

ROHLF, F. J., 1990. Rotational fit (Procrustes) methods. Pp. 227-236. In F. J. Rohlf & F. L. Bookstein (Eds.) *Proceedings of the Michigan Morphometrics*

Workshop. Special Publication No. 2, University of Michigan Museum of Zoology: Ann Arbor.

ROHLF, F. J. & F. L. BOOKSTEIN, 1987. A comment on shearing as a method for size correction. *Systematic Zoology,* 36: 356-367.

ROHLF, F. J. & D. SLICE, 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology,* 39: 40-59.

SEBER, G. A. F., 1985. *Multivariate Observations*. Wiley and Sons: New York.

SNEDECOR, G. W. & W. G. COCHRAN, 1967. *Statistical Methods*. 6th., ed. Iowa State Univ. Press: Ames.

SOKAL, R. R. & F. J. ROHLF, 1981. *Biometry*. 2nd. ed. W.H. Freeman and Co.: New York.

STEUNES, S. & L. F. MARCUS, 1991. A statistical study of species and sex differences in crania of Recent African and subfossil Madagascan Hippopotamids. In *Taxonomy and Paleobiology of the subfossil Madagascan pygmy hippopotami*. Uppsala University: Sweden.

THORPE, R., 1988. Multiple group principal component analysis and population differentiation. *Journal of Zoology,* 216: 37-40.

VOSS, R. S. L., L. F. MARCUS & P. ESCALANTE, 1990. Morphological evolution in muroid rodents. I. Conservative patterns of craniometric covariance and their ontogenetic basis in the neotropical genus *Zygodontomys*. *Evolution* 44: 1568-1587.

RELATIVE WARP ANALYSIS AND AN EXAMPLE OF ITS APPLICATION TO MOSQUITO WINGS

F. JAMES ROHLF

Department of Ecology and Evolution
State University of New York
Stony Brook, NY 11794-5245
ROHLF@SBBIOVM. BITNET

CONTENTS

ABSTRACT

This paper provides an alternative description of Bookstein's (1989, 1991) method of relative warps for the analysis of within-population morphometric variation for landmark data. The properties of the α parameter (Bookstein, 1991), that determines the relative weighting of the principal warps at different scales, is investigated. It is suggested that a value of $\alpha = 0$ is appropriate for taxonomic and exploratory studies where one has no *a priori* expectation that variation at a particular scale will be more important. In such cases it may also be useful to include information on affine differences among the specimens.

New techniques for the graphical representation of the results of a relative warp analysis are presented. The relationships between relative warp analysis and standard morphometric techniques such as canonical variates analysis, Fourier analysis, Procrustes analysis, and the analysis of coordinate data are described. Data on 18 landmarks from the wings of 8 species of *Anopheles* mosquitoes are used as an example to illustrate the methods.

The methods described in this paper are implemented by the TPSRW computer program for IBM PC compatible microcomputers.

INTRODUCTION

The method of relative warps is a technique developed by Bookstein (1989, 1991) for the analysis of within-population morphometric variation based on landmark data. Bookstein (1991) gives a detailed presentation of the mathematical basis for the method and furnishes examples of its application to cranial growth in rats and to the analysis of Apert's syndrome, a craniofacial anomaly in humans. The primary goal of the present paper is to provide an alternative description of this technique. It is hoped that a description from a somewhat different perspective will help make this important new method more accessible. Another purpose of this paper is to suggest several new graphical techniques for representing the results of a relative warp analysis.

The relationships between analyses of relative warps and standard multivariate and morphometric techniques such as canonical variates analysis, Fourier analysis, and Procrustes analysis are also discussed. The dependence of estimated distances between pairs of specimens on the parameters of the methods and the relationship to analyses of the original coordinate data are given special emphasis. In that respect this paper is a continuation of Rohlf (1986, 1990a, 1990b, and 1992).

This paper also serves as a more formal description of the TPSRW program than is practical to provide in its "README" file. This program (for IBM PC compatible microcomputers) is available via FTP over the Internet from SBBIOVM.SUNYSB.EDU and upon request from the author. Most of the computations and illustrations shown below were prepared using that program (the NTSYS-pc program was also used).

THE METHOD OF RELATIVE WARPS

Briefly, the method consists of fitting an interpolating function (the thin-plate spline of Bookstein, 1989) to the x, y-coordinates of the landmark for each specimen in a sample. Variation among the specimens within a sample is described in terms of variance in the parameters of the fitted functions. This is expressed relative to a *bending energy* matrix (see below) based on the coordinates of the landmarks of a reference configuration. The reference is often the mean configuration of

landmarks after some appropriate alignment of specimens. The relative warps are simply principal components vectors in this space and are used to describe the major trends in shape variation among specimens within a sample as deformations in shape (non-uniform shape variation).

Computation of relative warps

The computational steps described below are based in part on the algorithm given in section 7.6.2 of Bookstein (1991). Where appropriate, the correspondence to his step numbers is indicated. For simplicity of presentation, only the two-dimensional case is described below. The generalization to three dimensions is straight-forward (only the definition of the U function and the dimension of some of the matrices need to be changed). Those familar with canonical vectors analysis may find the steps easier to follow if they recognize the close analogy between canonical variates and relative warps and the use of the bending energy matrix as if it were a pooled within-groups variance-covariance matrix.

1. Let $\mathbf{X_i}$ denote the $2\times p$ matrix of the digitized x, y-coordinates of the p landmarks digitized for the i th specimen. We will use $\mathbf{X}$ to denote the $2n\times p$ matrix of coordinates for all n specimens in the sample. It will also be convenient to use $\mathbf{X_x}$ to refer to the $n\times p$ matrix of just the x-coordinates and $\mathbf{X_y}$ to refer to the corresponding matrix of the y-coordinates.
2. A reference configuration of landmarks must be obtained. The choice of a reference is important since the relative warps are relative to the eigenvectors of the bending energy matrix (see below) which is solely a function of the reference configuration.

   One approach is to use a reference specimen. This could represent an earlier developmental stage or a hypothetical common ancestor. More often it will simply be the mean location of the landmarks after the specimens have been aligned in some way. A simple method of alignment is to use Bookstein's (1986) method of shape coordinates. Bookstein (1991) uses shape coordinates to align each object relative to an *a priori* defined baseline. The location of each landmark in the consensus can be computed, for example, as the mean x, y-coordinates across the n aligned objects.

   Another approach is to use a superposition (Procrustes) method to construct a reference configuration. Sneath (1967) and Siegel and Benson (1982) describe methods for comparing pairs of organisms. Gower (1971) describes a generalization that allows many specimens to be compared simultaneously. The generalized affine resistant-fit method of Rohlf and Slice (1990) seems to be particularly effective in ignoring the effects of a few deviant landmarks and in achieving an intuitively appealing alignment that localizes the differences among specimens (unfortunately this method does not optimize any known goodness of fit criterion and its statistical properties are poorly known). This

method was used to construct a consensus configuration to use as a reference configuration in the examples given below. This consensus configuration is close the majority of specimens in the sample (after adjusting for their differing location, orientation, size, and shearing). The resulting $2 \times p$ *consensus configuration* matrix can be denoted $\mathbf{X}_c$. It may also be convenient to use the x, y-coordinates of the fitted objects in subsequent computations rather than the raw data especially if the original data were digitized with inconsistent alignments. If the objects differ greatly in size one may wish to scale them by their centroid size. Centroid size is the sum of squared distances between all pairs of landmarks or, equivalently, the sum of the squared distances of each landmark to the centroid of the specimen (Bookstein, 1991, pp. 93-94). This scaling can be done, for example, by dividing the coordinates for each object by the square root of average centroid size (centroid size divided by p), as done by Sneath (1971). Gower (1971) used the same normalization but without the division by p. This step corresponds to step 1 of Bookstein (1991).

3. Compute the bending energy matrix, $\mathbf{L}_p^{-1}$, for the consensus object. To do this one must first assemble the partitioned matrix

$$\mathbf{L} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{Q}^t & \mathbf{0} \end{bmatrix}, \tag{1}$$

where

$$\mathbf{P} = \begin{bmatrix} 0 & U(r_{12}) & U(r_{13}) & \cdots & U(r_{1p}) \\ U(r_{21}) & 0 & U(r_{23}) & \cdots & U(r_{2p}) \\ U(r_{31}) & U(r_{32}) & 0 & \cdots & U(r_{tp}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ U(r_{p1}) & U(r_{p2}) & U(r_{p3}) & \cdots & 0 \end{bmatrix}. \tag{2}$$

The U function is defined as

$$U(r_{ij}) = r_{ij}^2 \ln r_{ij}^2 , \tag{3}$$

where r_{ij}^2 is the square of the distance between landmarks i and j in the reference object. Note that the sign of U is as in Bookstein (1991) which is opposite from that in Bookstein (1989). The matrix

$$\mathbf{Q} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_p & y_p \end{bmatrix} \tag{4}$$

consists of a column vector of all ones followed by the x, y-coordinates of the reference object and **0** is a 3×3 matrix of all zeros.

The bending energy matrix, $\mathbf{L}_p^{-1}$, is the upper-left p×p block of the inverse of matrix **L**. The product $\mathbf{X}_t \mathbf{L}_p^{-1}$ yields the coefficients for the non-affine part of the thin-plate spline that transforms the coordinates of landmarks in the reference configuration into those of the "target" specimen, i. The affine coefficients are given by the product $\mathbf{X}_t \mathbf{L}_q^{-1}$, where $\mathbf{L}_q^{-1}$ is the upper-right p×3 block of the inverse of matrix **L**.

4. Decompose the bending energy matrix as $\mathbf{L}_p^{-1} = \mathbf{E}\,\Lambda\,\mathbf{E}^t$, where Λ is a p×p diagonal matrix of eigenvalues and **E** is the p×p matrix of eigenvectors (the columns of which correspond to the normalized eigenvectors and the rows to the landmarks). Bookstein (1989) calls these eigenvectors the *principal warps*. The magnitudes of the eigenvalues are inversely related to scale. Large eigenvalues correspond to eigenvectors that describe small-scale features (deformations of landmarks that are close together). Small eigenvalues correspond to eigenvectors that describe large-scale features. At least three of the eigenvalues will be equal to zero since they correspond to the affine components that are of infinite scale (translation, rotation, and dilatation). These zero eigenvalues and their corresponding eigenvectors can be deleted to reduce Λ to a p-3×p-3 matrix and **E** to a p×p-3 matrix. The dimensions corresponding to the affine components are not lost since they can still be computed from the $\mathbf{L}_q^{-1}$ matrix described above (see further discussion below). These operations correspond to steps 2 and 3 of Bookstein (1991).
5. Compute a weight matrix, **W**, as a scaled projection of the x and y-coordinates of the deviations of the n objects from the reference object onto the principal warps.

$$\mathbf{W} = [\ \mathbf{W}_x \mid \mathbf{W}_y\ ], \tag{5}$$

where

$$\mathbf{W} = \frac{1}{\sqrt{n}}\ \mathbf{V}\ (\mathbf{I}_2 \otimes \mathbf{E}\ \Lambda^{-\alpha/2}), \tag{6}$$

$$\mathbf{V} = [\ \mathbf{V}_x \mid \mathbf{V}_y\ ], \tag{7}$$

and

$$\mathbf{V}_x = \mathbf{X}_x - \mathbf{1}_n \otimes\ [0/1]\ \mathbf{X}_c\ . \tag{8}$$

$$\mathbf{V}_y = \mathbf{X}_y - \mathbf{1}_n \otimes\ [0/1]\ \mathbf{X}_c\ . \tag{9}$$

The symbol ⊗ is used above to denote a direct (Kronecker, tensor) product of two matrices. The $\mathbf{V}_x$ matrix is an n×p matrix of the x-coordinates of the differences between the n objects and the reference object ($\mathbf{1}_n$ is a column vector of n ones). $\mathbf{V}_y$ is the corresponding matrix of y-coordinates. Alternatively, one could uses deviations from the sample means as in

Bookstein (1991)—rather than use deviations from the reference object. This step assumes that the objects have been aligned in some reasonable way. The coefficient $1/\sqrt{n}$ in Equation 6 does not appear explicitly in Bookstein (1991). It is implied since a variance-covariance matrix (with a division by n) was used.

The elements of the $n\times2(p\text{-}3)$ weight matrix **W** describe each specimen as a linear combination of the principal warps. Since the Λ matrix is of rank p-3, Equation (6) represents a projection of the p-dimensional space of variation at each landmark (separately for each coordinate) onto a p-3 dimensional subspace. What this projection leaves behind is variation among the specimens with respect to translation, rotation, and uniform shape change.

Bookstein (1991) has suggested the introduction of the parameter α in Equation (6). If $\alpha > 0$ then only those principal warps that have eigenvalues greater than zero can be used (in order to avoid having to divide by zero). This is why those dimensions were deleted in the previous step. A value of $\alpha = 1$ yields the relative warp analysis as described by Bookstein (1989) in which the principal warps are weighted inversely by the square roots of their eigenvalues. This means that large-scale variation (variation among specimens in the relative positions of widely separated landmarks) is given more weight than small-scale variation (variation in the relative positions of landmarks that are close together).

The ith row of **W** corresponds to the linear combination of the normalized principal warps that would yield the non-linear component of the thin-plate spline that transforms the reference object's configuration of landmarks into those of the ith object. The coefficients for these thin-plate splines are given as the rows of

$$\mathbf{N} = \mathbf{W}\sqrt{n}\,(\mathbf{I}_2 \otimes \Lambda^{1+\alpha/2}\,\mathbf{E}^t). \tag{10}$$

This matrix is of dimension $n\times2(p\text{-}3)$. Using Equation (6), this can be simplified to

$$\begin{aligned}\mathbf{N} &= \mathbf{V}(\mathbf{I}_2 \otimes \mathbf{E}\Lambda\mathbf{E}^t)\\ &= \mathbf{V}(\mathbf{I}_2 \otimes \mathbf{L}_p^{-1})\end{aligned} \tag{11}$$

6. A singular-value decomposition (Eckart and Young, 1936, Jöreskog, et al., 1976) is then performed to yield the following factorization of the weight matrix:

$$\mathbf{W} = \mathbf{S}\,\mathbf{D}\,\mathbf{R}^t\,, \tag{12}$$

where **S** is a matrix of normalized scores with its rows corresponding to the n objects and the columns corresponding to the min (n-1, 2 (p-3)) relative warps with singular values>0, **D** is a diagonal matrix of singular values (it is of the same dimension as the columns of **S**), and **R** is a matrix whose columns correspond

to the relative warps and the rows correspond to the weighted principal warps. The first p-3 rows of **R** pertain to the x-coordinates and the remaining pertain to the y-coordinates for each landmark. The columns of **S** and **R** are normalized to length 1. All of the relative warps do not need to be retained just those that account for an appreciable proportion of the total variance among the specimens (energy normalized if $\alpha > 0$). The relative warp scores are uncorrelated since the columns of S are orthogonal.

Bookstein (1991) used a different approach in his steps 6 and 7 and only considered the case with $\alpha = 1$. He computed the eigenvalues and eigenvectors of the variance-covariance matrix of the x, y-coordinates of each point scaled in terms of inversely weighted principal warps (each principal warp with $\lambda_i > 0$ was divided by $\sqrt{\lambda_i}$. This procedure yields the same matrix, **R**, of relative warps but his eigenvalue matrix is equal to $\mathbf{D}^2$. The matrix of normalized scores was computed by him using the relationship $\mathbf{S} = \mathbf{W}\,\mathbf{R}\,\mathbf{D}^{-1}$.

7. The matrix of relative warps should be expressed in terms of the original x, y-coordinate system rather than in terms of the principal warps. These *relative warp loadings* (Bookstein, 1991) can be computed as

$$\mathbf{R}' = (\mathbf{I}_2 \otimes \mathbf{E}\Lambda^{-\alpha/2})\mathbf{R} \tag{13}$$

Each column of **R'** can be represented as a displacement vector at each landmark in the reference object (see below). The rows correspond to the p pairs of x, y-coordinates and the columns correspond to the relative warps. This corresponds to step 8 of Bookstein (1991). The relative warps can also be modeled as thin-plate splines (see below).

It is difficult to visualize the major morphometric components of variation among the specimens by an examination of just the numerical results described above (*e.g.*, by studying the entries in the **S** and **R'** matrices). Fortunately, there are a number of graphical displays that allow one to visualize the statistical results in terms of the 2 or 3-dimensional space that the specimens were digitized in rather than only in terms of multivariate vector spaces. Some suggestions and example are provided in the following sections.

Graphical presentations of the results of a relative warp analysis

Several suggestions on ways to display the results of a relative warp analysis are described below. The first three methods are especially useful since they allow one to superimpose the results on plots of the specimens themselves. This has the advantage that it keeps one's attention focused on the geometry of the configurations of landmarks in the 2 or 3-dimensional space of the organism. This will make it

easier to visualize the results in terms of changes in the shape of the organisms and should make the results easier to interpret biologically. The last two graphics are conventional displays of multivariate vector spaces. They are useful for looking for clusters and other patterns in the relationships among the specimens.

1. The relative warp loadings can be shown as displacement vectors at each landmark on the reference specimen. Bookstein (1991) uses the elements of the normalized relative warp loadings matrix, **R'**. Since these are not in the same units as the x, y-coordinates of the landmarks, the vectors must be scaled arbitrarily to make them of convenient length for plotting in the same space as the digitized reference specimen. Alternatively, one could use the scaled loading matrix, **R' D**, so that their lengths would be proportional to the square roots of the variance per unit bending energy for each relative warp (arbitrary scaling is still required, however). Figures 3 and 6, below, are examples.

   One problem with such plots is that one is tempted to interpret them as indicating how each landmark would be displaced by the effect of each relative warp. The vectors are related to the coefficients of a thin-plate spline (see below for details). The actual displacement at a location x, y is a weighted sum of $r_i^2 \ln r_i^2$ values (where r_i is the distance from x, y to landmark i). Loading vectors will match those of the actual displacements only when a relative warp is closely aligned with a principal warp (*i.e.*, when the corresponding column of **R** contains essentially all zeros except for a particular principal warp).
2. One can show all of the original objects superimposed on the reference object as is done in the various types of Procrustes or superposition analyses (*e.g.*, Rohlf and Slice, 1990). In addition, the displacements at each landmark can be shown as vectors as described above. Since the affine component of the thin-plate splines are not provided by Equation (11), an inverse transformation based on the affine components of a thin-plate spline can be used. The inverse transformation is computed as follows for specimen i:

$$[\ \mathbf{A_0} \mid \mathbf{A_1}\ ] = \mathbf{X}_i\, \mathbf{L}_Q^{-1} \tag{14}$$

   where $\mathbf{L}_Q^{-1}$ is the upper right $p \times 3$ block of the inverse of matrix **L** (Equation (1)), $\mathbf{A_0}$ is a column vector of dimension 2 containing the displacements for x and y, and $\mathbf{A_1}$ is a 2×2 shear matrix. Specimen i can be superimposed on the reference specimen by transforming its coordinates for each landmark as:

$$\mathbf{X}_i' = \mathbf{A_1^{-1}}\ (\ \mathbf{X}_i\ \mathbf{A_0 1}_p^t)\ , \tag{15}$$

   where $\mathbf{1}_p$ is a column vector of p 1s.

   When α is greater than 0, landmarks that are closer together have more of an influence on the parameters of the affine transformation than those landmarks that are far apart. The among-specimen scatter of each landmark

around its position in the reference specimen represents the variation that is described by the relative warps. Note also that even though the specimens may all have been carefully aligned (*e.g.*, by a generalized affine resistant fit analysis, Rohlf and Slice, 1990), this step will still adjust the alignment of the specimens since it is using a different criterion for fit. In order to see the variation that the relative warps are attempting to describe one must use the "minimum-energy" superimpostion option in the TPSRW program.

3. Each relative warp can be plotted as a deformation of the space of the reference configuration of landmarks. This can be shown by computing a thin-plate spline for each relative warp. The non-affine coefficients for these thin-plate splines can be computed as:

$$\mathbf{N'} = \sqrt{n}\ (\ \mathbf{I}_2 \otimes \mathbf{E}\ \Lambda^{1+\alpha/2}\ )\ \mathbf{R}\ \mathbf{D} \qquad (16)$$

This gives the coefficients for a unit change in a relative warp score from that of the reference object.

These deformations can be shown as animated displays in which the reference object is deformed as a thin-plate spline in a positive direction along a selected relative warp axis and then in a negative direction. These displays are very useful for visualizing the integrated overall change in shape implied by a set of displacement vectors at each landmark.

One can also show displacement vectors at each landmark by connecting the position of each landmark in the reference configuration with its new location in the transformed space. The pattern of vectors is usually similar to that of the relative warp loadings. Figures 4 and 7, below, are examples.

4. The coefficients of the thin-plate splines can be regressed on age, a measure of size, or other variables of interest. Rather than simply reporting the numerical results, one can make an animated display (like a movie) using the thin-plate splines to show how the reference configuration would be expected to change as a function of changes in the independent variable. The TPS-REGR program performs these operations if necessary, one could log-transform the independent variable or use non-linear regression methods. This would, for example, allow one to discover regions of the organism that show allometry without having to decide in advance which linear distance measurements should be regressed on size. For more on allometry and its regionalization, see Bookstein (1991). For studies involving allometry the use of $\alpha = 1$ is recommended since the effects of allometry tend to be at larger scales.
5. The scaled scores,

$$\mathbf{S'} = \mathbf{S}\ \mathbf{D} \qquad (17)$$

can be plotted against each other for the first few relative warps to provide an ordination of the specimens. One expects specimens with similar configurations of landmarks (after removing any differences that a thin-plate spline treats as

an affine transformation) to be close together in the ordination space. Thus one can search for clusters of similar specimens, look for trends, correlate with exogenous variables, etc. The position of the ith object indicates the importance of each relative warp in determining the thin-plate spline that would transform the reference object into the ith specimen. Figures 5 and 8, below, are examples.

Note that these ordinations only take into account differences that represent non-affine deformations and these differences are weighted inversely by their bending energies if $\alpha > 0$. The affine components may contain useful information about shape variation if the initial positions of the specimens have been aligned using a criterion that yields a better alignment than that provided by the affine component of a thin-plate spline. For one approach to producing a scatter of the affine components per se, see Bookstein (1991, Sec. 7.2). These are conceptually a first pair of relative warps those for $\lambda = 0$, the pair at the largest possible scale.

6. A biplot of the weight matrix, **W**, can be made by superimposing a plot of the columns of **R** (corresponding to the principal warps) on the plot of scaled scores, **S'**, described above. However, one would rather express the biplot in terms of the thin-plate splines since they are expressed in terms of the coordinates for each landmark. The matrix, **N**, of the coefficients of the thin-plate splines for each object can be expressed as the product $\mathbf{S'R''}$, where

$$\mathbf{R''} = \sqrt{n}\,(\,\mathbf{I}_2 \otimes \mathbf{E}\,\Lambda^{1+\alpha/2}\,)\,\mathbf{R} \qquad (18)$$

The matrix **N'** differs from **R''** only by the multiplication by matrix **D**. Matrices **S'** and **R''** can be plotted simultaneously to yield a biplot of the thin-plate splines for each specimen as a function of the relative warps.

Choice of metrics

Bookstein (1991 and personal communication) has suggested the introduction of the parameter α in the exponent of Λ in computing the weight matrix, **W**, defined in Equation (6). A value of $\alpha = 1$ corresponds to the relative warp analysis described by Bookstein (1989, 1991). This value results in variation among specimens in those principal warps that have relatively small bending energies (corresponding to large-scale features in the reference configuration) being weighted more heavily than variation in those principal warps with larger bending energies (corresponding to relatively small-scale features). On the other hand, the principal warps at the largest scale (the affine components) are ignored since they have bending energies equal to zero. The decomposition in Equation (12) gives the directions of maximum variance among specimens relative to the bending energy matrix just as canonical variates analysis (*e.g.*, Krzanowski, 1988, or Reyment, 1991) gives directions of maximum variance among groups relative to the within-population variance-covariance matrix.

Other values of α can also be used to give different relative weightings to the principal warps. The case of $\alpha = 0$ (suggested by Bookstein, 1991, p. 368) seems to be of particular interest for exploratory studies. This value gives all of the principal warps the same weight. Thus the analysis is no longer *relative* to bending energy—even though the principal warps are used as the basis vectors for the space. In this space the Cartesian distance between specimens is the affine-free Procrustes distance between specimens. In terms of the original coordinate data, it corresponds to the Cartesian distance between the x and y-coordinates of a pair of specimens after differences explainable by affine transformations (translation, rotation, and uniform stretching) have been removed. Note however that the affine components removed are as defined by the method of thin-plate splines. Other methods may give a different partitioning between the affine and non-linear differences between specimens.

If $\alpha = 0$ then one can simplify the matrix of relative warp loadings in Equation (13) to

$$\mathbf{R}' = (\mathbf{I}_2 \otimes \mathbf{E})\,\mathbf{R} \tag{19}$$

Since the columns of **E** are orthonormal, multiplication by **E** corresponds to a rigid rotation of a linear vector space. However, the bending energy matrix is not of full rank since at least 3 eigenvalues are equal to 0.

It is important to note that, despite the use of non-linear functions such as thin-plate splines and the U function, the relative warp loadings and the relative warp scores are just linear combinations of the original x and y-coordinates of the specimens. Different choices of reference configurations or of α simply correspond to different rotations and weightings of the original x and y-coordinates. On the other hand, the principal warps which serve as basis vectors for the space and the values of the weights assigned to them are non-linear functions of the coordinates of the landmarks in the reference configuration. They vary in complex ways as one varies the reference configuration. The use of different reference configurations and of α effects the expression of the final conclusions obtained. Unfortunately, the choices must be somewhat arbitrary since thin-plate spline functions do not correspond to a biological model for developmental or evolutionary shape change. They simply represent a method for capturing such changes (in contrast, for example, to the approach of Ackerly, 1990).

Analysis of affine variation

The above account is an incomplete description of the variation among specimens within a sample since it only considers variation that can be explained in terms of deformations. The components that span the space of affine differences (translation, rotation, and uniform shape change) among individuals have been

explicitly removed from the analysis by ignoring the last 3 eigenvectors of the bending energy matrix. While a separate procedure (*e.g.*, Bookstein, 1991, Section 7.2) can be used to describe such differences, I believe it will often be useful to analyze both types of variation simultaneously. It would be interesting to know, for example, that those individuals that have high scores on relative warp 1 are also smaller and more elongate than the average. Several approaches are suggested below.

An obvious technique is to correlate rotation angle, centroid size, uniform factor score, strain cross parameters, and perhaps even translation values (see below) for each specimen with the relative warp scores. This is somewhat inefficient since these added variables may be partially redundant. This does not give a very elegant overall analysis.

An alternative approach is to append additional variables to the **W** matrix before performing the singular-value decomposition described in step 6 of Section 2.1. To be comparable to the existing elements, the coefficients should correspond to the affine coefficients of the thin-plate splines that transform the reference configuration into that of each specimen. These can be computed as $\mathbf{X}_i \mathbf{L}_q^{-1}$ which was set aside in the computations of the relative warps. In this way the resulting component axes will summarize both uniform and non-uniform shape variation. One must take into account the fact that these coefficients are not orthogonal to the principal warps and are not in the same units. A solution is simply to retain the last 3 eigenvectors of the bending energy matrix. The matrix **E** will then remain a $p \times p$ matrix. If $\alpha = 0$ then the Λ matrix can be ignored. Otherwise it must be modified so that the λ_i are taken as equal to unity (rather than 0) for the last three eigenvalues. These last three normalized eigenvectors correspond to that part of the affine variation that is orthogonal to the principal warps.

The consequence of retaining all of the eigenvectors is mostly interesting for the $\alpha = 0$ case. The distances between pairs of specimens will then be the same whether based on the **V** matrix (the original matrix of coordinates of the specimens) or the **S'** matrix (the matrix of scaled relative warp scores). This means that the method of relative warps does not change one's perception of the relative distances between among the specimens. Thus projection onto the principal components axes based on the original coordinate data will be the same as the relative warp scores. This is because the **V** matrix differs from the **W** matrix only by its multiplication on the *right* by the orthonormal matrix $\mathbf{I}_2 \otimes \mathbf{E}$. In this case relative warp analysis simply provides an interpolating function that allows one to describe and reconstruct morphometric variation in terms of a convenient continuous function. This is analogous to the fitting of elliptic Fourier coefficients (Kuhl and Giardina, 1982, Rohlf, 1986), parametric cubic splines (Evans *et al.*, 1985, Rohlf, 1990a), and other functions of outlines rather than computing the empirical eigenshape functions (Lohmann, 1983, Lohmann and Schweitzer, 1990). If the affine components are not included then the $p \times (p\text{-}3)$ matrix **E** projects the data into a

space of lower dimension and some information will be lost if these dimensions are ignored.

AN EXAMPLE OF RELATIVE WARP ANALYSIS APPLIED TO MOSQUITO WINGS

The dataset

The techniques described above were applied to the first $n = 8$ species (all in the genus *Anopheles*) in the mosquito wing dataset used in Rohlf and Slice (1990). These data serve as a convenient test dataset (these data are included with the TPSRW program and the reader is encouraged to try to duplicate the results given here and to try further experiments). There are $p = 18$ landmarks corresponding to points at which wing veins either branch or intersect the margin of the wing. Figure 7 of Rohlf and Slice, 1990, gives the standard nomenclature for the veins and also shows the positions of the landmarks that were used in this study. The names of the species used and their code numbers used in this study are shown in Figure 1. Also shown are the positions of the 18 landmarks. The landmarks are connected with solid lines to represent the approximate topology of the wing veins (the actual veins are not just straight lines).

A reference configuration was constructed using generalized affine resistant fit analysis (Rohlf and Slice, 1990). The reference configuration is shown in Figure 1 with the veins shown as dotted lines. In most cases the fit to the reference is very good—especially near the tip of the wing. Species 5 shows a particularly poor fit near the base of the wing. The positions of the landmarks in the reference configuration are also indicated by the origins of the various vectors shown in many of the figures presented below. The scatter of each specimen's landmarks around the position of the landmark in the reference can be seen in Figures 3 and 6 below. As discussed in Rohlf and Slice (1990) for the complete dataset, the scatter around the landmarks near the tip of the wing (landmarks 3 through 9) is much less than at most of the other landmarks-especially those at the base of the wing (landmarks 1, 12, and 13) and landmarks at the leading and trailing edges of the wing (landmarks 2 and 11).

There are $p–3 = 15$ principal warps with eigenvalues greater than zero that can be extracted from the bending energy matrix. The relative magnitudes of the eigenvalues (from 124.1407 down to 0.0698) is a function of the spatial arrangement of the landmarks. For the mosquito wing data the second largest eigenvalue is much smaller than the largest eigenvalue because the average wing is very elongated. The principal warps corresponding to the largest-scale features (smallest non-zero eigenvalues) of the reference configuration are illustrated in Figure 2. Note that they are shown, arbitrarily, as displacements to both the x and y-coordinates and that the magnitude for each principal warp was scaled

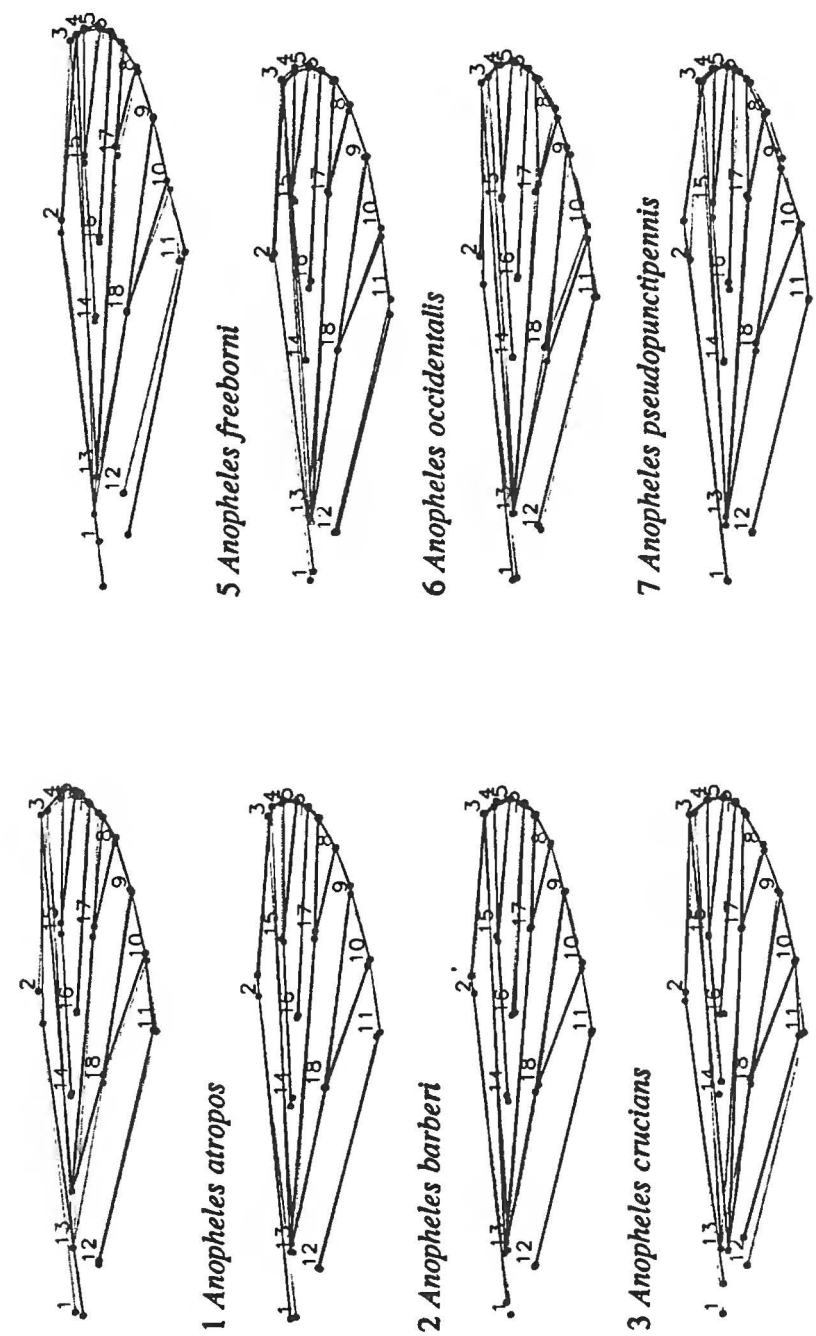
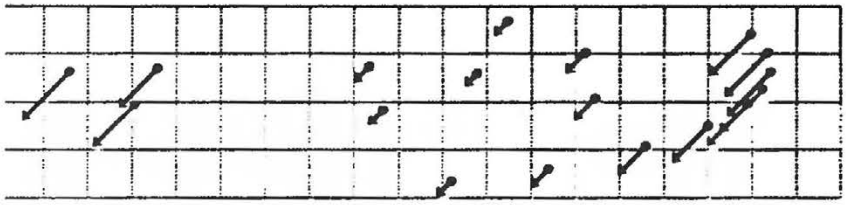
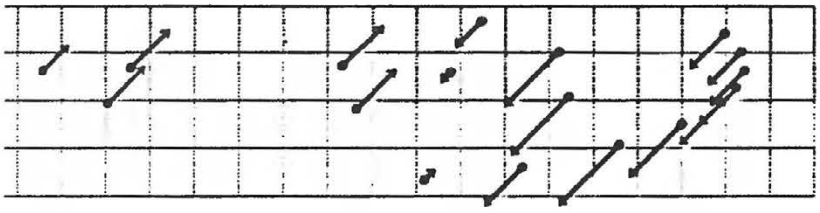


Fig. 1

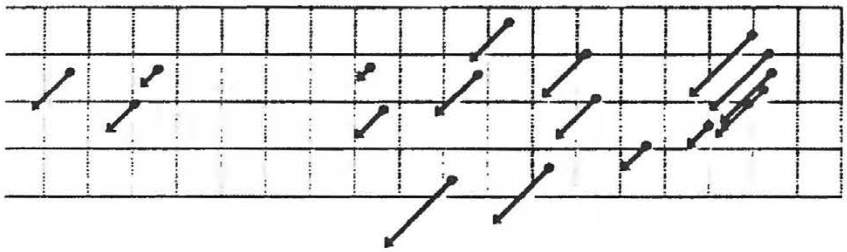
Plots of the 8 species of Anopheles mosquitoes used in the present study superimposed on the reference configuration using affine generalized resistant-fit analysis. The reference configuration is shown with dotted lines



Principal warp 15



Principal warp 14



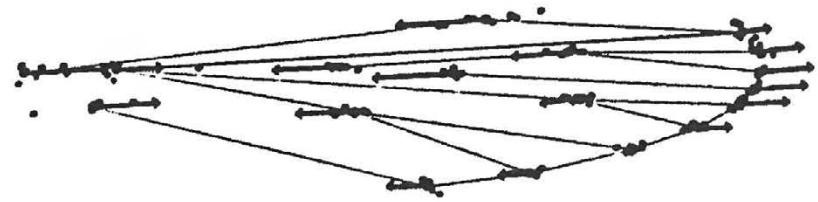
Principal warp 13

Fig. 2
Plots of the last 3 principal warps shown as displacement vectors equally for both x and y-coordinates

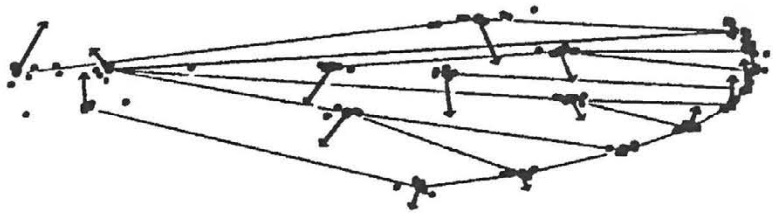
arbitrarily. The relative lengths of the vectors at different landmarks indicates the relative weighting of each landmark for a given principal warp.

Relative warps analysis, with $\alpha = 1$

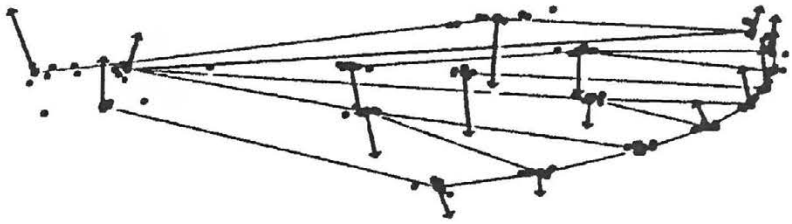
There are 7 relative warps that can have eigenvalues greater than zero (0.406, 0.108, 0.080, 0.063, 0.047, 0.020, and 0.016). These are the square roots of the variance per unit bending energy. The relative warp loadings (the **R'** for the relative warps with the largest eigenvalues are illustrated as vectors at each landmark in Figure 3. The vectors point in directions of maximum variance of the joint scatter



Relative warp 1



Relative warp 2



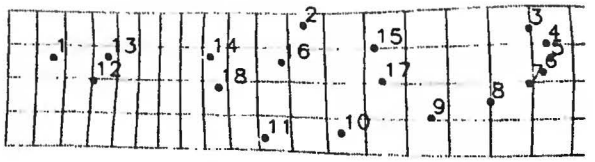
Relative warp 3

Fig. 3

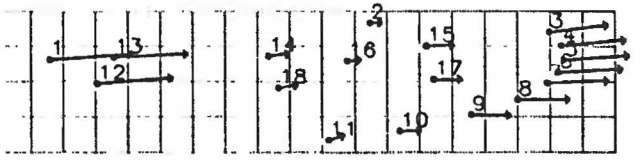
Plots of the relative warp loadings (using $\alpha = 1$) and the scatter of the 8 wings superimposed on the refernce configuration based on an affine generalized resistant-fit

among the 8 wings relative to bending energy. The absolute magnitudes of the vectors are arbitrarily scaled for each warp. One can see that the vectors indicate only large-scale variation. Relative warp 1 indicates an expansion of the region near the tip relative to a compression near the base. Relative warp 2 indicates a straightening of the leading edge of the wing by movement of the central landmarks towards the trailing edge of the wing. Relative warp 3 appears to indicate a similar deformation but with the landmarks at the tip of the wing moving forward.

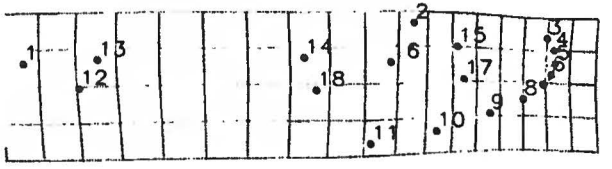
The relative warps can also be illustrated as thin-plate splines as shown in Figure 4. This figure shows the deformations implied by positive and negative displacements along the first two relative warps. While the vectors for relative warp 1 are similar to those shown in Figure 3, the vectors for the second relative



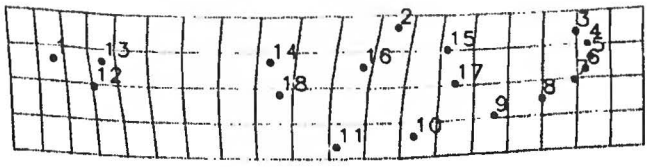
Relative warp 1 + deformation



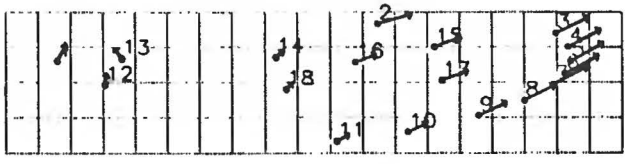
Relative warp 1 + deformation shown as vectors



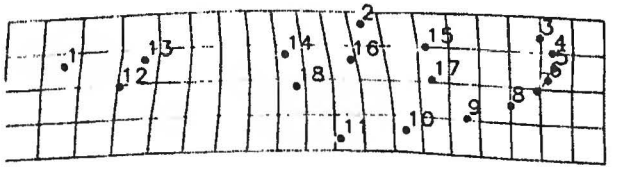
Relative warp 1 − deformation



Relative warp 2 + deformation



Relative warp 2 + deformation shown as vectors



Relative warp 2 − deformation

Fig. 4

Plots of the first two relative warps (based on α = 1) shown both as thin-plate splines for positive and negative displacements along the first and second relative warp aces and as vectors showing the net movement of the landmarks

warp seem rather different. These figures show what a specimen would look like if its relative warp score were at an extreme position along one of the relative warp axes and zero on all others. These figures are analogous to those of figure 9 of Rohlf and Archie (1984) for a principal components analysis based on Fourier coefficients. Rohlf (1992) compares ordination analyses based on the coefficients of various functions fitted to outline data.

Relative warp 1 is mostly a function of a single principal warp (this can be seen in matrix **R**, not shown). Its highest coefficients are -0.992 on the *x*-coordinates and -0.054 on the *y*-coordinates of principal warp 15 (the largest-scale principal warp). The next largest coefficient is -0.057 for the *x*-coordinates of warp 12.

The second relative warp has its highest coefficients for the *y*-coordinates of principal warp 15 (-0.542), *x*-coordinates of principal warps 14 (-0.527), and 12 (0.486). The next highest coefficient is for the *x*-coordinates for principal warp 7 (0.220). The fact that the most important relative warps are most closely aligned with the last few principal warps is expected since they have been given much larger weights. For example, the weight given to the last principal warp is much larger than that given to the first, $\sqrt{\lambda_1/\lambda_{15}} = \sqrt{124.1407/0.06985} = 42.2$. It would require a very large amount of among-specimen variation in principal warp 1 in order for it to be represented among the first few relative warps.

Figure 5 shows a scatter-plot of the relative warp scores (projections of the 8 species onto the relative warp axes) for first two relative warps. Most of the variation among species (relative to bending energy) is along the first relative warp axis with species 5 at the extreme left and species 1 at the extreme right. From Figure 4 one expects (and easily sees) that the largest differences between species 1 and 5 are that the central landmarks (14 to 18) are closer to the tip of the wing in species 1 and closer to the base in species 5. Species 3 is at the top of the plot and species 6 is at the bottom of the small cloud of points (species 1 is further down but at the extreem right). Figure 4 implies that landmarks 1, 12, and 13 at the base of the wing should be relatively more anterior in species 3 (making the leading edge of the wing seem less forward) in species 3 than in species 6. It is harder to see this in species 1 since the effect of relative warp

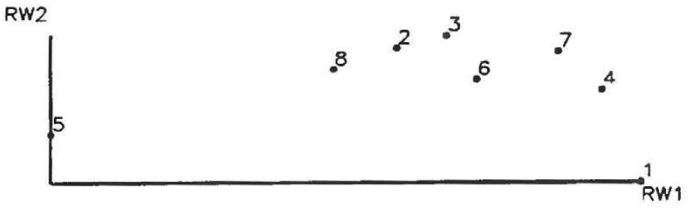


Fig. 5

Plot of the 8 species of mosquitoes with respect to their relative warp scores for the firs 2 relative warps

is so strong. These predictions can be checked against the original specimens as illustrated in Figure 1. Species 5, indicated to be somewhat of an outlier in Figure 5, does seem to show the largest residuals from the reference configuration. The first two relative warps explain 84.0% of the variance (relative to bending energy). Relative warp analysis strongly weights differences between the specimens in the largest-scale features. In a subjective examination of the wings shown in Figure1 one considers all features (but with an unknown weighting).

Relative warps analysis, with $\alpha = 0$

The results obtained using the Procrustes metric, $\alpha = 0$, were quite different from those reported above. The eigenvalues are 0.165, 0.108, 0.102, 0.057, 0.047, 0.033, and 0.026. The first relative warp does not dominate as much as was found with $\alpha = 1$. The relative warp loadings (the **R'** for the relative warps with the largest

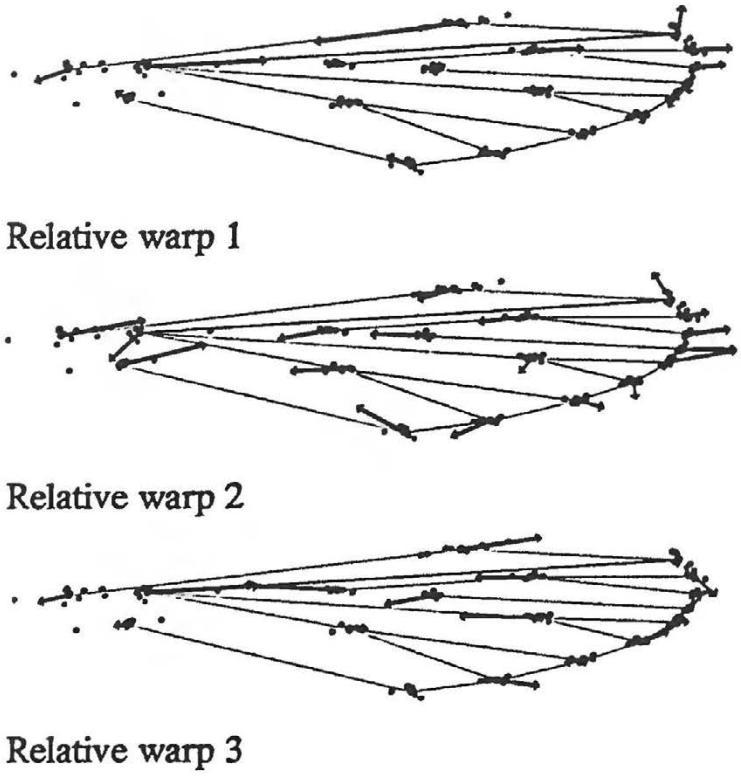


Fig. 6

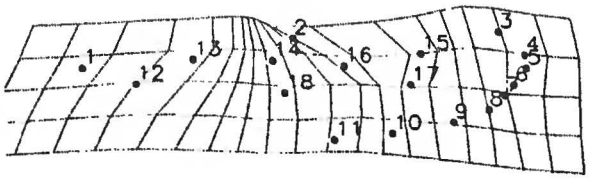
Plots of the relative warp loadings (using $\alpha = 0$) and the scatter of the 8 wings superimposed on the reference configuration based on an affine generalized resistant-fit

eigenvalues) are illustrated as vectors at each landmark in Figure 6. The vectors point in directions of maximum variance of the joint scatter among the 8 wings—but not relative to bending energy since $\alpha = 0$. The absolute magnitude of the vectors are arbitrarily scaled as before. As expected, the patterns of displacements are more complex and localized than those shown in Figure 3 in which the larger scale features were heavily weighted. The longest vectors for relative warp 1 imply displacements of landmark 2 towards the base of the wing and landmark 13 towards the tip. This seems to match well with the fact that there is a large amount of scatter at those landmarks and in a direction parallel to the vectors. Relative warp 2 indicates a more complex result with landmarks 1 and 12 being displaced away from the base and landmark 13 towards the base (the net effect being that these landmarks should become closer together). The central landmarks 2, 10, 11, and 18 displaced towards the base of the wing and some of the landmarks at the tip displaced away from the base (indicating an expansion in that region). Relative warp 3 indicates a compression of the region between landmarks 13 and 14.

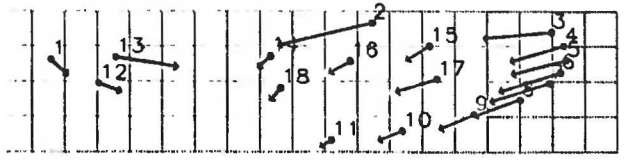
The relative warps are illustrated as thin-plate splines in Figure 7. This figure shows the deformations implied by positive and negative displacements along the first two relative warps. The vectors for relative warp 1 imply a somewhat different pattern of deformation to those shown in Figure 6 (Figures 4 and 7 are more similar than one might have expected from a comparison of Figures 3 and 6). The most apparent deformation is the displacement of landmark 2 and the landmarks at the tip of the wing towards the base relative to the displacement of most of the other landmarks towards the tip of the wing. The second relative warp indicates the movement of landmark 13 towards the base, a general expansion near the center of the wing and an outward displacement of the landmarks at the tip of the wing. The illustrations of positive and negative displacements show what a specimen would look like if its relative warp scores were at an extreme position along one of the relative warp axes and zero on all others.

An important difference from the results with $\alpha = 1$ is that now many of the principal warps contribute to the first few relative warps. The largest contribution to relative warp 1 is the x-coordinates for principal warp 7 (-0.543). The other contributors are the x-coordinates for principal warps 12 (-0.501), 4 (-0.376) and 15 (-0.331). The major contributors to relative warp 2 are the x-coordinates of principal warps 15 (-0.812), 8 (-0.247), and 12 (0.206). Only x-coordinates are involved in the first few relative warps in contrast to the results for $\alpha = 1$ where the second relative warp showed a strong displacement in the y-direction.

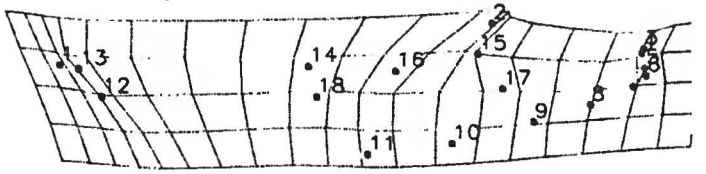
Figure 8 shows an ordination scatter-plot of the relative warp scores (projections of the 8 species onto the relative warp axes) for first two relative warps. The first relative warp axis has somewhat more variance than the second. Species 1 is at the extreme right of the axis 1, as in Figure 5, but the distribution of the other points is different. Species 8 is now at the extreme left and the other points are more spread out. Species 6 is now at the bottom of the plot.



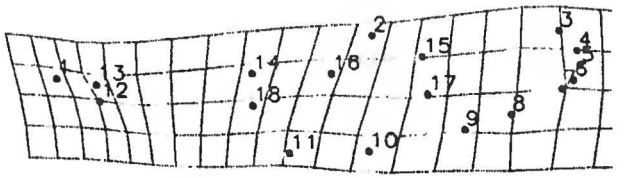
Relative warp 1 + deformation



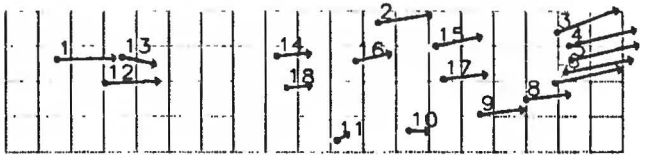
Relative warp 1 + deformation shown as vectors



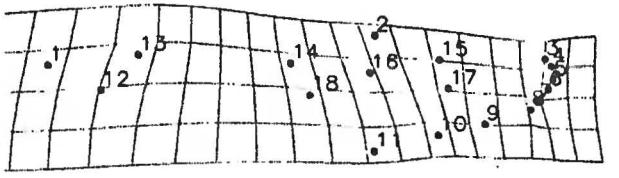
Relative warp 1 − deformation



Relative warp 2 + deformation



Relative warp 2 + deformation shown as vectors



Relative warp 2 − deformation

Fig. 7

Plots of the first two relative warps (based on α = 0) shown both as thin-plate splines for positive and negative displacements along the first and second relative warp axes and as vectors showing the net movement of the landmarks

One would expect from Figure 7, that the most obvious difference between species 1 and species 8 would be that landmark 2 should be displaced towards the base of the wing relative to the other landmarks in species 1 (there should also be an expansion of landmarks 1, 12, and 13 near the base). Species 4 is now at the top of the plot and species 6 is at the bottom. Figure 7 implies that landmarks 1 and 12 at the base should be closer to 13 and that the tip of the wing expanded (landmarks 3 to 7 should be more displaced away from 8, 9, 15, and 17) in species 4 than in species 5. These predictions match what one can see in Figure 1 (one does not expect a perfect match, however, since the first two relative warps only explain 50.8% of the non-affine variance). Note that the amount of variance explained relative to bending energy (as found using $\alpha = 1$) is not directly comparable to this value since the former is normalized by bending energy. Using all of the relative warps for this dataset, the matrix correlation between distances among specimens based on the scaled score matrix and distance based on the original coordinates of the specimens is 0.879 (for $\alpha = 1$ the correlation is only 0.841). If the effects of the affine components were retained then the correlation would, of course, rise to 1.0 (the ordination along the first two axes remains very similar to that shown in Figure 8).

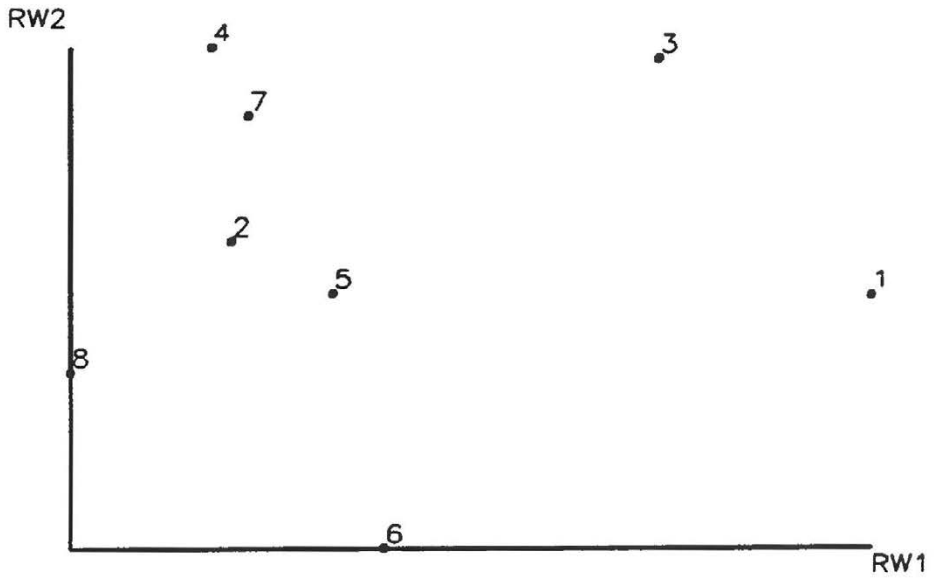


Fig. 8
Plot of the 8 species of mosquitoes with respect to their relative warp scores for the first 2 relative warps based on $\alpha = 0$.

DISCUSSION

The description of the method of relative warps and the example of its application given above raises a number of important questions.

1. What value should be used for the α parameter? The effect of using a value greater than zero is to give more emphasis to the larger-scale features. If one were to use a value less than zero it would result in more weight being given to the small-scale features. The different choices can make appreciable differences in the final results. For allometric growth studies or other applications where large scale differences are expected a value of $\alpha = 1$ is likely to be the most useful. In many exploratory studies (such as in taxonomy) it may not be clear which features should be given more weight than others and thus a value of $\alpha = 0$ which gives an equal weighting is likely to be the most appropriate. This latter choice corresponds to using an affine-free Procrustes distance as the measure of morphometric distance between specimens. In a study such as that of Weber (1992) one may wish to try a value closer to $\alpha = -1$. In many studies one should try a range of values of a in order to search for interesting patterns at different scales.
2. Is bending energy a useful parameter in morphometrics? This quantity is based on the physical properties of the thin metal sheets that are the basis of the thin-plate splines. The fact that it takes more energy to make a smaller-scale deformation in a sheet of metal than it does to make a larger scale-one does not seem to be necessarily an appropriate model for the amount of developmental or evolutionary "effort" it might take to achieve a certain deformation of a configuration of landmarks. If the energy parameter does not seem meaningful, then it does not seem appropriate to use it to weight the principal warps. This is another reason for using $\alpha = 0$. On the other hand, one can view bending energy as just a convenient index for scale without attaching biological significance to the parameter itself.
3. Should the three eigenvectors corresponding to the affine components of the bending energy matrix be retained with the other principal warps (at least for the $\alpha = 0$ case)? Including translation and rotation in the analysis assumes that the specimens have been aligned in some meaningful way and are not just artifacts of the digitization process. Including these components results in an analysis that describes the total within-population variation not just non-uniform shape differences. This seems to a convenient way to detect covariation between these two different kinds of shape differences.
4. Do different choices of reference configurations make much difference in the final results? The answer is clear for the case of $\alpha = 0$. Different choices simply result in different orthogonal rotations of the basis vectors of the relative warp space and thus will yield identical results. However, when $\alpha \neq 0$ the results will differ depending upon the choice of reference configuration. This

is because different features are given different weights depending upon their proximity in the reference configuration. It seems unlikely that there should be a single best solution for obtaining a reference configuration. If one is studying variation among specimens sampled from a homogeneous population then some sort of average configuration (such as constructed by generalized resistant-fit analysis) seems reasonable. If, on the other hand, one is studying variation along a developmental or evolutionary sequence then an estimate of a configuration for an early or a primitive stage may be more appropriate. Fortunately, the results seem fairly stable for small changes in the positions of the landmarks in the reference configuration. As a result different choices of reference configurations often make very little difference in the final result.

5. What are the relative advantages of the method of Procrustes superpositions versus the computation of relative warps? I believe that the method to be preferred will depend upon the type of variation one expects to find. If one expects differences in only a very small proportion of the landmarks to be displaced relative to the others (the "Pinocchio effect" of Chapman, 1990), then Procrustes methods provide a direct simple solution with an appropriate graphical display. When variation is not well localized, Procrustes plots are less effective. Procrustes plots show the relative levels of variation at different landmarks but it is difficult to appreciate the pattern of covariation between the displacements at different landmarks. The method of relative warps displays such covariation very effectively. For example, one can see in Figure 7 that as landmark 2 varies towards the base of the wing, landmark 13 moves away from the base and landmark 3 moves forward in a direction orthogonal to the other displacements.

   Another advantage of the use of the method of relative warps is that it is possible to use thin-plate splines to construct hypothetical configurations to represent points in the parameter space. This is very useful since it allows one to visualize means of clusters, endpoints of axes, etc. even if the functions themselves are not based on a biological model (as in applications of Fourier analysis in morphometrics). While the Procrustes method can be thought of as just a special case ($\alpha = 0$) of relative warp analysis, Reyment's (1991) conclusion that Procrustes superposition methods have been made largely obsolete seems somewhat premature. Relative warp analysis uses a particular method (thin-plate splines) for aligning specimens and for separating affine from non-affine variation that may not be appropriate in all cases. Even though resistant-fit methods may have been used to determine the reference configuration and the initial alignment of the specimens, the variation actually analyzed by relative warp analysis is that which is present after an alignment based on thin-plate splines.

While additional work remains to be done, it is already clear that the method of relative warps is a flexible and powerful technique. It should become part of the standard morphometric tool kit for the analysis of landmark data.

ACKNOWLEDGMENTS

The patience of Fred Bookstein in clarifying many of the details of the method of relative warps is greatly appreciated. Les Marcus and Dennis Slice also read versions of the ms. and contributed important suggestions.

This work was supported, in part, by a grant (no. BSR-89-18630) from the Systematic Biology Program of the National Science Foundation.

This paper is contribution number 817 from Graduate Studies in Ecology and Evolution, State University of New York at Stony Brook.

REFERENCES

ACKERLY, S. C., 1990. Using growth functions to identify homologous landmarks on mollusc shells. Pp. 339-344 In Rohlf, F. J. & F. L. Bookstein (eds.) *Proceedings of the Michigan Morphometrics Workshop*. Museum of Zoology special publication no. 2, University of Michigan: Ann Arbor.

BOOKSTEIN, F. L., 1986. Size and shape spaces for landmark data in two dimensions (with discussion and rejoinder). *Statistical Science*, 1: 181-242.

BOOKSTEIN, F. L., 1989. Principal warps: Thin-plate splines and the decomposition of deformations. *I.E.E.E. Transactions on Pattern Analysis and Machine Intelligence*, 11: 567-585.

BOOKSTEIN, F. L., 1990. Four metrics for image variation. In *Proceedings of the XI international conference on information processing in medical imaging* (Ortendahl, D. & J. Llacer, eds.) Alan R. Liss, Inc, New York.

BOOKSTEIN, F. L., 1991. *Morphometric tools for landmark data*. Cambridge University Press. 435 pp.

CHAPMAN, R. E., 1990. Conventional Procrustes approaches. Pp. 251-267. In Rohlf, F. J. & F. L. Bookstein (eds.) *Proceedings of the Michigan Morphometrics Workshop*. Museum of Zoology special publication no. 2, University of Michigan: Ann Arbor.

ECKART, C. & G. YOUNG., 1936. The approximation of one matrix by another of lower rank. *Psychometrika*, 1: 211-218.

EVANS, D. G., P. N. SCHWEITZER, & M. S. HANNA., 1985. Parametric cubic splines and geological shape descriptions. *Mathematical Geology*, 17: 611-624.

GOWER, J. C., 1971. Statistical methods of comparing different multivariate analyses of the same data. Pp. 138-149. In *Mathematics in the archaeological and historical sciences* (Hodson, F. R., D. G. Kendall & P. Tautu, eds.). Edinburgh University Press, Edinburgh.

JÖRESKOG, K. G., J. E. KLOVAN, AND R. A. REYMENT., 1976. *Geological factor analysis*. Elsevier: Amsterdam. 178 pp.

KRZANOWSKI, W. J., 1988. *Principles of multivariate analysis: a user's perspective*. Oxford University Press, Oxford. 563 pp.

KUHL, F. P. & C. R. GIARDINA, 1982. Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, 18: 236-258.

LOHMANN, G. P., 1983. Eigenshape analysis of microfossils: A general morphometric procedure for describing changes in shape. *Mathematical Geology*, 15(6): 659-672.

LOHMANN, G. P. & P. N. SCHWEITZER, 1990. On eigenshape analysis. Pp. 147-166. In Rohlf, F. J. and F. L. Bookstein (eds.) *Proceedings of the Michigan Morphometrics Workshop*. Museum of Zoology special publication no. 2, University of Michigan: Ann Arbor.

REYMENT, R. A., 1991. *Multidimensional paleobiology*. Pergamon Press: New York, 377 pp.

ROHLF, F. J., 1986. Relationships among eigenshape analysis, Fourier analysis, and analysis of coordinates. *Mathematical Geology*, 18: 845-854.

ROHLF, F. J., 1990a. Fitting curves to outlines. Pp. 167-177. In Rohlf, F. J. and F. L. Bookstein (eds.) *Proceedings of the Michigan Morphometrics Workshop*. Museum of Zoology special publication no. 2, University of Michigan: Ann Arbor.

ROHLF, F. J., 1990b. Morphometrics. *Annual Review of Ecology and Systematics*, 21: 299-316.

ROHLF, F. J., 1992. The analysis of shape variation using ordinations of fitted functions. In *Ordinations in the study of morphology, evolution and systematics of insects: applications and quantitative genetic rationales*. (Sorensen, J. T., ed.) Elsevier, Amsterdam. In press.

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

ROHLF, F. J. & D. SLICE., 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, 39: 40-59.

SIEGEL, A. F. & R. H. BENSON, 1982. A robust comparison of biological shapes. *Biometrics*, 38: 341-350.

SNEATH, P. H. A., 1967. Trend-surface analysis of transformation grids. *Journal of Zoology*, 151(1): 65-122.

WEBER, K. E., 1992. How small are the smallest selectable domains of form? *Genetics*, 130: 345-353.

THE FRACTAL ANALYSIS OF SHAPE

DENNIS E. SLICE

Ecology and Evolution
State University of New York
Stony Brook, NY 11794
DSLICE@SBBIOVM.BITNET

CONTENTS

ABSTRACT

Many structures of interest to biologists are of such complexity that they cannot be adequately characterized by simple measurements. Fractal analysis provides a method for the quantification of such complexity by means of the fractal dimension, D. This statistic summarizes the changes in estimates of length, area, or other measures with changes in the precision of the measurement. This, in turn, is directly related to the form and degree of the complexity of the material being considered. In the first section, this paper presents an overview of mathematical fractals, their properties, and methods for the estimation of D. Applications of the fractal analysis in a number of biological fields are reviewed. The second section illustrates the approach through the analysis of leaf outlines from several species of the genus *Acer* (maple trees). Problems involved in the practical application and interpretation of fractal analysis are discussed.

INTRODUCTION

Many structures of interest to biologists are of such complexity they cannot be adequately characterized by simple measurements or by the landmark-based techniques discussed elsewhere in this volume. This chapter will discuss a technique, fractal analysis, that directly uses this complexity to construct a summary measure called the fractal dimension, D. This measure can be treated like any other descriptor and used to investigate environmental, evolutionary, or other factors that might influence or be influenced by the complexity of a particular structure.

The first section of this chapter provides a brief introduction to fractals that describes what they are, some of their unique and interesting properties, and how these properties can be used to describe the complexity of real objects. It also presents an overview of some of the ways in which fractal analysis has been used to answer diverse questions in biology. The second section concerns the study of shape variation in outlines of leaves from trees in the genus *Acer*. This part includes an application of fractal analysis that points out some critical, but somewhat subtle, problems of such an analysis.

A single chapter can provide no more than a general introduction to fractal analysis. This should be enough to give some sense of the potential and limitations of the technique and allow readers to assess its suitability to their own research interests. There are several texts that can be recommended for further reading. *The Fractal Geometry of Nature* (Mandelbrot, 1983) provides a summary and synthesis of earlier work by the author who is most responsible for bringing fractals to the attention of the broad audience they have today. Unfortunately, Mandelbrot's expansive knowledge and free-form style of writing make this text somewhat difficult for many readers. A quite clear presentation of the mathematics and ideas behind fractals can be found in *The Science of Fractal Images* (Peitgen & Saupe, 1988), and *Fractals* by J. Feder (1988) provides an exceptional discussion of the basic concepts of fractal analysis with emphasis on its application to the study of real world phenomena. *Fractals Everywhere* by Michael Barnsley (1988) is also a most useful volume.

FRACTALS AND FRACTAL ANALYSIS

What is a fractal?

Most of us have been exposed to fractals at least in the form of renderings of exquisitely complex mathematical structures like Mandelbrot and Julia sets. But what is it that makes them special and qualifies them to be called fractals? Mandelbrot (1983) defines a fractal as "a set for which the Hausdorff-Besicovitch dimension strictly exceeds the topological dimension." This definition, while mathematically rigorous, is perhaps too formal and restrictive for a general discussion. A less stringent form suggested by Mandelbrot (see Feder, 1988) is that "a fractal is a shape made of parts similar to the whole in some way." It is the latter definition that we shall adopt and we will be primarily concerned with the "way" in which the parts are "similar to the whole." In particular, when the overall form of a structure is repeated at smaller scales within itself, the structure is said to be *self-similar,* and it is this self-similarity, real or assumed, that is the basis for the fractal analysis taken up later.

Figure 1 illustrates two fractal shapes that meet the self-similarity criteria in different ways. Figure 1A is a rendering of the triadic Koch curve along with the components used in its construction. Shapes like this are generated by starting with a simple, initial component, called an *initiator,* and a more complicated structure, called a *generator,* made up of scaled-down copies of the initiator. The original initiator is first replaced by the generator. The copies of the initiator in the resulting structure are then replaced with appropriately scaled versions of the generator. This process is, in theory, repeated an infinite number of times to produce a curve that at scales below that of the original initiator is composed entirely of small, exact copies of itself. In the case of the triadic Koch curve, the initiator is simply a straight line segment and the generator is a line segment of identical length that has had the middle third replaced with two line segments one-third the length of the original. In cases such as this, where identical copies of the whole can be found in the parts, the curve is said to have *exact* self-similarity. Figure 1B shows a simulated fractal coastline which represents a second type of self-similar curve. Here the exact form of the entire curve is found nowhere in the small-scale parts. Instead, it is its overall complexity that is retained. Such a relationship is termed *statistical* self-similarity.

The existence of self-similarity, assumed present at all magnifications, leads to the property that fundamentally distinguishes fractal curves from those that are non-fractal, or Euclidean. As one examines a fractal curve in greater and greater detail there is no tendency for it to "smooth out". On the other hand, Euclidean curves, no matter how seemingly complex, will eventually tend to smooth out into straight lines at infinitesimal scales of observation. This leads to the result

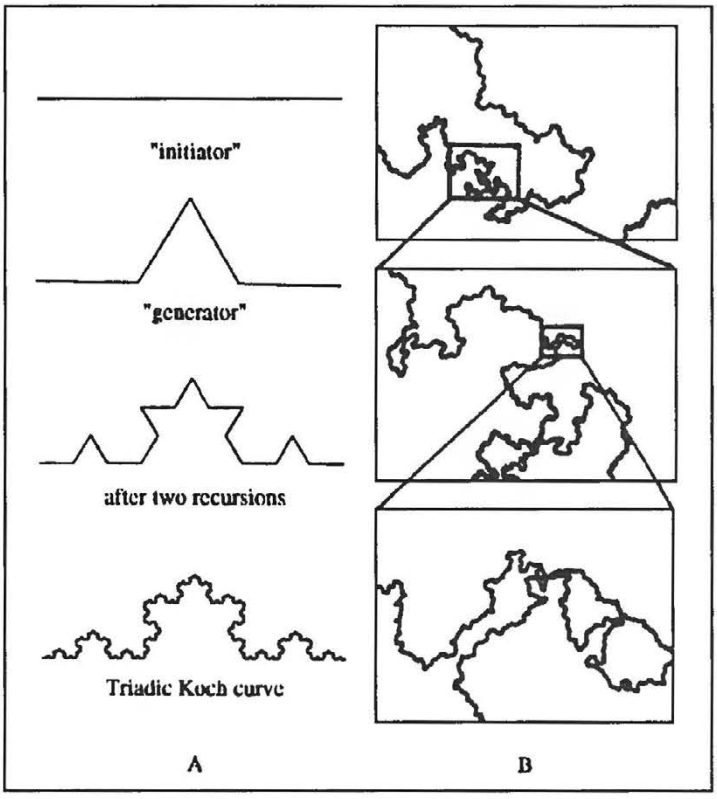


Fig. 1

A) Fractal curve generated by the repeated replacement of "initiators" by "generators". Smaller parts of the curve are exact copies of the whole leading to exact self-similarity. B) Repeated magnification of a statistically self-similar fractal coastline. Small-scale parts retain the complexity, but not the exact shape, of the whole. (after Voss, 1988)

that as one increases measurement precision, the measured length of Euclidean curves will eventually converge on a single value while length measurements of fractal curves will diverge true fractal curves have an infinite length. The rate at which length estimates of fractal curves diverge is directly related to the complexity of the curve and will be used to characterize that complexity in the form of the fractal dimension, D.

One-dimensional curves and outlines are not the only structures that can exhibit self-similarity. Two-dimensional surfaces can also have the same relative roughness at all scales while their Euclidean counterparts smooth out into a collection of vanishingly small planar patches. In this case, it is the area that converges for Euclidean surfaces and diverges for fractals as the measurement precision increases. Three dimensional structures, too, can have distributions that are inhomogeneous at all scales.

Another potentially self-similar pattern found in biology is that of branching structures. Such patterns can be generated in a manner similar to that used in the construction of Koch curve by replacing the initiator by a symmetrically or asymmetrically forked generator and subsequently applying the generator only to the newly formed branches. The model can be extended to transport structures by specifying relative branch diameters at each splitting. Basically, anything that requires efficient communication with all parts of a two- or three-dimensional object would be a candidate for fractal branching. Examples include plant branch and root systems, bifurcating bronchial structures, and vascular or neural networks. Horsfield (1990) and West & Goldberger (1987) provide a more detailed discussion of branching patterns and Glenny *et al.* (1991) give a general review of fractals with a good discussion of branching patterns emphasizing applications in physiology.

What is fractal analysis?

We have seen that fractals have the unique characteristic that as one increases the resolution with which they are examined, estimates of their "size", e.g. length, area, etc., tend to diverge to infinity instead of converging on a particular value. The rate of this divergence is related to their complexity and quantified by the measure D, the fractal dimension. To see how we can estimate the value of D for real data we start by measuring some familiar Euclidean shapes.

First, consider a line segment. We could determine its length by stepping over it with a divider or ruler of a particular length, r, and counting the number of steps, N(r), required to cover the segment. The product of the number of steps and the length of a step would then be our measure of the length, L(r), of the segment based on the step size, r, $L(r) = N(r)r$. If r were equal to the total length of the line segment, we would require but single step to traverse the segment, N(1)=1. Letting r be one-half of the total length would require two steps, N(1/2)=2, and for r=1/3, N(1/3)=3. One can see the general equation describing this pattern is $N(r)=r^{-1}$.

To extend this procedure to determine the area of two-dimensional regions we could use squares of a particular area and count the number of such squares necessary to cover the region. Since the area of a square is just its edge-length, r, squared, the formula for calculating the area would be $A(r)=N(r)r^2$. If the region being measured is itself a square and we write r as a proportion of its edge-length then: for r=1, N(1)=1; for r=1/2, N(1/2)=4; and for r=1/3, N(1/3)=9. The general equation for this relationship is $N(r)=r^{-2}$. Similarly, if we apply this approach to measuring the volume of a cube by filling it with smaller cubes the relationship between number and size of measuring device is found to be $N(r)=r^{-3}$.

All of the above cases show a consistent relationship between "size" and the scale of the measuring device given by the equation

$$N(r)=r^{-D} \quad (1)$$

where D is referred to as the similarity dimension. In the case of the simple Euclidean shapes, the similarity dimension is the same as the topological dimension. For self-similar fractals, the similarity dimension exceeds the topological dimension and equals the Hausdorff-Besicovitch dimension (Feder, 1988).

If we use the same procedure to measure the Koch curve in Figure 1A, we see that for r=1, relative to the length of the initiator, the number of steps required to traverse the curve is 1 as with the line segment. However, if we use a ruler with r = 1/3, four steps are required and the measured length is increased to N(1/3)1/3 = 4/3. This is because the smaller ruler is able to include the additional length due to the large, triangular bump in the middle of the curve that could not be resolved with the larger ruler. Using a still smaller ruler, say with r set to 1/9, would result in an even longer length measurement, 16/9, due to the inclusion of even smaller bumps resulting from the second application of the generator. If Equation 1 holds, and it does, then D cannot be the same as the topological dimension of the Koch curve, which is one. By taking logarithms of both sides of Equation 1 and rearranging we can get the equation necessary to determine the value of D. That formula is

$$D= -\ln(N(r)) / \ln(r) \quad (2)$$

For the Koch curve using r=1/3 we find D=-ln(4)/ln(1/3)=1.26. With r=1/9 we have D=–ln(16)/ln(9)=1.26. In fact, since Equation 1 applies, for any choice of r we would find D=1.26. The Koch curve is a self-similar fractal, and as such, its similarly dimension exceeds its topological dimension — D is the fractal dimension.

Most natural phenomena are not likely to have the neat, orderly structure of the triadic Koch curve, but instead, more closely resemble the statistically self-similar curve of Figure 1B. This does not present any major problems. First, if the relationship in Equation 1 holds then we need only measure the length of the curve with some ruler of relative length r and use Equation 2 to determine D as with the Koch curve. However, we cannot be certain that Equation 1 holds at all length scales. Also, we do not know r relative to the "size" of the curve. To address the latter problem, we choose some absolute ruler size, λ, and scale it by the, as yet unknown, maximum length of the curve, L_{max}. The value for r can then be written as $r=\lambda/L_{max}$. Using the available information we can write the relationship between the estimated length of our curve and the selected ruler size as

$$\hat{L} = N(\lambda)\lambda = r^{-D}\lambda = \frac{L_{max}^{D}}{\lambda^{D-1}}.$$

Some simple algebra gives us the more useable form

$$\ln(\hat{L}) = D \ln\left[L_{max}\right] + (1-D)\ln(\lambda) \qquad (3)$$

which resembles the slope-intercept equation of a line, $Y = a + bX$. We can, therefore, analyze an arbitrary curve by getting a number of estimates of the length of the curve using a variety of ruler lengths and performing a linear regression of the log-transformed length estimates onto log ruler length. The value of D can then be estimated as 1–b. Also note that the unknown L_{max} is found in the constant term and can be estimated as $e^{a/D}$.

The procedure described above can be used to estimate D for any continuous curve and will be used later in this chapter to examine variation in the complexity of leaf outlines. Similar methods can be developed that, assuming the curve lies within a plane, cover the plane with a grid of scale r and count the number of boxes through which the curve passes. In this case, $N(r) \sim r^{-D}$. This technique, called *box-counting*, also allows for the analysis of not just outlines, but also of point or areal distributions.

Finally, there is an important, but often overlooked, point that should be emphasized, and that is the distinction between the mathematical study of fractals and the application of fractal techniques to the study of natural phenomena. The fractal sets studied by mathematicians are generally the products of deterministic equations or of carefully planned recursive constructions. These entities generally manifest a fractal character at all scales of observation. Real-world data need not be so well-behaved. Practical considerations such as digitizer resolution limit the range over which objects can be observed, and even over a limited range there is no guarantee that biological processes will produce the scaling behaviour assumed in the estimation procedure. Natural objects are no more guaranteed to be fractals than any two variables are guaranteed to have a linear relationship. It is only assumed that growth and developmental processes produce structures sufficiently fractal-like that D captures some important aspect of the complexity of the structure. The assumed linear relationship between log length estimate and log step size must be checked and no extrapolations beyond the range of scales examined can safely be made. Fractal analysis may provide an adequate description of, or insight into, a particular phenomenon, but if not, other tools must be sought.

Applications of fractal analysis in biological sciences

The potential of fractal analysis for the study of shape is evident by the diversity of problems to which it has been applied. It has been used to investigate the structure of components of organisms, whole body forms of individual and colonial organisms, the relationship of organisms to the structure of their environment, and more. Table 1 shows a sampling of D-values published by various authors. Some of these examples are discussed below and should serve to stimulate ideas for other applications.

Table 1

A sampling of published D values. Values represent means for different data sets, ranges within a data set, or representative values. See source for details

DATA	D	SOURCE
HABITAT:		
Aleutian Islands	1.19 - 1.66	Pennycuick and Kline, 1986
Canadian Shield lake shorelines	1.10 - 1.64	Kent and Wong, 1982
coral reefs	1.05 - 1.15	Bradbury *et al.*, 1984
branches and twigs	1.28 - 1.79	Morse *et al.*, 1985
WHOLE ORGANISM:		
Streptomyces and *Ashbya*		
mycelia outline	1.34 - 1.52	Obert *et al.*, 1990
mass	1.36 - 1.52	"
Serratia marcescens colonies	1.4 - 1.6	Matsuyama *et al.*, 1989
Trichoderma viride colonies	1.4 - 2.0	Ritz & Crawford, 1990
mite wanderings	1.092 - 1.117	Dicke & Burrough, 1988
WITHIN ORGANISM:		
leaf shapes	1.02 - 1.22	Vlcek & Cheung, 1985
plant root systems	1.48 - 1.58	Tatsumi *et al.*, 1989
sutures		
white-tailed deer skulls	1.19 - 1.65	Long, 1985
ammonite shells	1.20 - 1.53	"
blood flow distribution		
cardiac (misc. spp.)	1.17 - 1.21	Glenny *et al.*, 1991
pulmonary (dogs)	1.07 - 1.12	"
mammograms of healthy breasts	2.22 - 2.50	Caldwell *et al.*, 1990
OTHER:		
taxonomic groupings	1.10 - 2.14	Burlando, 1990

Morse *et al.* (1985) represents one of the earliest applications of fractals analysis to the study of how the structural complexity of the environment can influence the abundance and distribution of organisms. They use the box-counting method to quantify the space-filling nature of a variety of plants and conclude that smaller

animals would be exposed to relatively greater spatial resources due to the fractal structure of vegetation. Using 1.5 as an average estimate for the fractal dimension of vegetation and incorporating considerations of metabolic rates and populations densities, they estimate how body size should be distributed for arthropod communities. Available data are consistent with their predictions. On a larger scale, Pennycuick Klein (1986) use the concept of fractally structured habitats to adjust bald eagle nesting density estimates on two differently structured islands. Kent Wong (1982) incorporate the fractal dimensions of lakes into estimates of littoral zone extent, and Phillips (1985) and Palmer (1988) consider fractal properties of the spatial structure of vegetation patterns and environmental gradients respectively, much of which was anticipated by Loehle (1983).

Caddy Stamatopoulos (1990) have recently presented an interesting synthesis of fractal analysis with traditional fisheries population biology by investigating the relationship between mortality curves and habitat complexity in crevice-dwelling organisms. Using growth and mortality data they attempt to deduce what would be a desirable distribution of crevice sizes in the environment. Alternatively, they anticipate the biological productivity of an arbitrarily chosen distribution of crevices. They find habitats with large D values should lead to higher juvenile survivorship, but be unable to support a large crop of adult organisms since crevice space becomes limiting for the numerous juveniles. This implies that less complex habitats may produce the same or more adults with less "wastage" of juveniles. These results are consistent with mangrove swamps and grass beds serving as important nursery areas while sustaining relatively few adults. The authors also propose an interesting experimental apparatus and designs for "sampling" crevice availability and identifying habitat bottlenecks.

The structure of a more abstract space is explored by Burlando (1990) who uses the size-frequency distribution of taxa based on the number of included subtaxa. He finds marine groups have a higher fractal dimension than continental ones and suggests this indicates marine environments somehow allow for the development of greater biological diversity. Other fractal phenomena above the level of oraganism morphology are discussed by Frontier (1987).

Some organisms lend themselves to the fractal analysis of their entire form. Obert *et al.* (1990) use box-counting methods to analyze both the outline and mass distribution of colonies of two microbial species, *Streptomyces griseus* and *Ashbya gossypii*. They find the fractal model fits the development of the outline and mass of both organisms well and that the fractal dimension increases during growth. Using the divider method, Matsuyama *et al.* (1989) are also able to quantify fractal patterns in the growth of *Serratia marcescens* colonies. They further determine that the inability to produce certain exolipids reduced the fractal structure of colonies and renewed fractal growth could be initiated by experimental application of these substances. Ritz Crawford (1990) relate the fractal growth patterns and changes in the complexity with growth of colonies of *Trichoderma viride* with foraging strategies for the exploitation of substrates with patchy and uniform food distributions.

Smaller parts and structures of an organism are likely candidates for fractal analysis. Leaf shapes are a good example and will be discussed in the next section. The first use of D to quantify differences in leaf morphology is due to Vlcek & Cheung (1986). They use the divider method to calculate D for leaves from eight species of trees. The leaf shapes range from relatively simple shapes of American basswood *(Tilia americana)* to the more complicated White oak *(Quercus alba)*. They find significant differences between three groups of species, but are unable to distinguish individual species. Tatsumi *et al.* (1989) use the box-counting method to analyze root networks of common agricultural plants such as the garden pea *(Pisum sativum)* and common millet *(Panicum miliaceum)*. In zoology, Long (1985) considers the use of fractal techniques to quantify complexity of sutures in deer skulls *(Odocoileus virginianus)* and shells of various species of extinct ammonites.

There are two recent medical applications of fractal analysis using two dimensional data. These suggest methods that might be of more general use in the biological sciences. Caldwell *et al.* (1990) classify mammograms by their fractal dimension. Mammograms are generally put into one of four categories called "Wolfe grades" based on the relative distribution of fat and visible ducts. There is a suggestion that some grades may be at greater risk of developing breast cancer. The authors find that trained radiologists have about an 85% agreement in classifying mammograms. Using density as the third dimension and different size grids to measure area and estimate the fractal dimension of the surface, they find classifications based on D to have an 84% agreement with the radiologists. This suggests an automated approach using D might be as effective as human interpretation and also admits the possibility of a continuous measure of mammogram structure that might provide a better assessment of cancer risk. Similarly, Lynch *et al.* (1991) use fractal "signatures" to identify differences in the texture of arthritic knee joints.

AN EXAMPLE

Lobed, cleft, parted, divided, erose, and undulate are just some of the terms botanists use to describe the shape of a leaf or its margin. Even with so many verbal descriptors natural variation can blur the distinction between shape classifications, and within a classification a considerable degree of variation remains unquantified. Yet, differences in the pattern and complexity of leaf outline can reflect taxonomic relationships, wide-scale geographic trends, or adaptation to local and microenvironmental conditions of importance to the organism and the scientist. Leaf outlines may not be obviously self-similar, but fractal analysis produces estimates of D that may effectively summarize important differences in shape and complexity. In this section I will demonstrate the use of fractal analysis to quantify variation in the shape of leaf outlines from several species in the genus *Acer* (maple trees).

Materials and methods

Eight trees were chosen from a larger study on leaf shape and physiology by Jessica Gurevitch. These included one individual of *Acer mono* (an East Asian maple), two *A. lobelii* (native to Northern Italy), two *A. saccharum* (North American sugar maple), two *A. palmatum* (varieties from Korea and Japan), and *A. japonicum* (a Japanese species). All trees were in Harvard University's Arnold Arboretum and had experienced the same climatic regime during growth.

Eighteen to forty-eight leaves were collected from various parts of the canopy of each tree. The leaves were gathered from locations selected to reflect the range of canopy microenvironments (sun and shade, northern and southern exposure, and different heights within the canopy) but were selected arbitrarily within locations. For the purpose of this analysis, all leaves collected from a tree will be treated as a random sample. Immediately after collection, the fresh leaves were photocopied and the copies checked for distortion by overlaying the original leaf on the copy — no differences were detected.

A computer imaging system was used to capture video images of the leaf copies. Images were enlarged to fill as much as possible of the image field which consisted of 512x480 picture elements (pixels). The *x,y*-coordinates of the outlines were then collected using the automatic edge detection capabilities of IMAGE (Rohlf & Slice, 1990).

The outlines were processed by FRACTAL-D ver 1.00 (Slice, 1989), a program that estimates the fractal dimension of an outline using the divider method and a set of user supplied step-lengths. The program randomly selects an initial reference point on an outline and moves from point to point along the outline until the straight-line distance to the reference point meets or exceeds that specified by the current step-length. The number of steps is increased by one and the point at which the step-length was exceeded becomes the new reference point. This is repeated until the entire outline, up to the initial reference point, is traversed. The length estimate is then made by multiplying the number of steps by the current step-length. The process is repeated for each specified step-length, and the fractal dimension estimated as 1-b, where b is the slope obtained from the regression of log length estimate onto log step size.

Each outline was enlarged to the same area. Fifty step-lengths from 0 to 403 pixels were used. The maximum step length was chosen to ensure that even the largest step lengths produced nonzero length estimates for every outline. The remaining step-lengths were selected so that after being log-transformed they were evenly spaced between 0 and ln(403). Five different starting points on each outline were used to determine the fractal dimension of each leaf.

The above procedure produced 1,270 estimates of D for 254 leaves from eight trees of five species. The GLM procedure in the Statistical Analysis System, SAS® (SAS Institute Inc., 1985), was used to carry out a nested analysis of variance

to determine the effects of species, tree-within-species, and leaf-within-tree on the fractal dimension of the leaf outlines. The SAS® procedure VARCOMP was used to estimate the distribution of variability in D below the species level.

A naive analysis

Representative outlines of each species are shown in Figure 2. One can see a progression of increasing complexity overlaid on a basic five-lobed pattern from the fairly simple *Acer mono* through the intermediately complex *A. saccharum* and ending with *A. japonicum* with its highly serrate margins and additional lobing. These differences are reflected in the mean D values (Table 2) and in the analysis of variance table (Table 3) where differences in D between species are highly significant ($p<0.0074$). Differences among trees within a species and among leaves within a tree are even more highly significant ($p<<0.001$ in both cases).

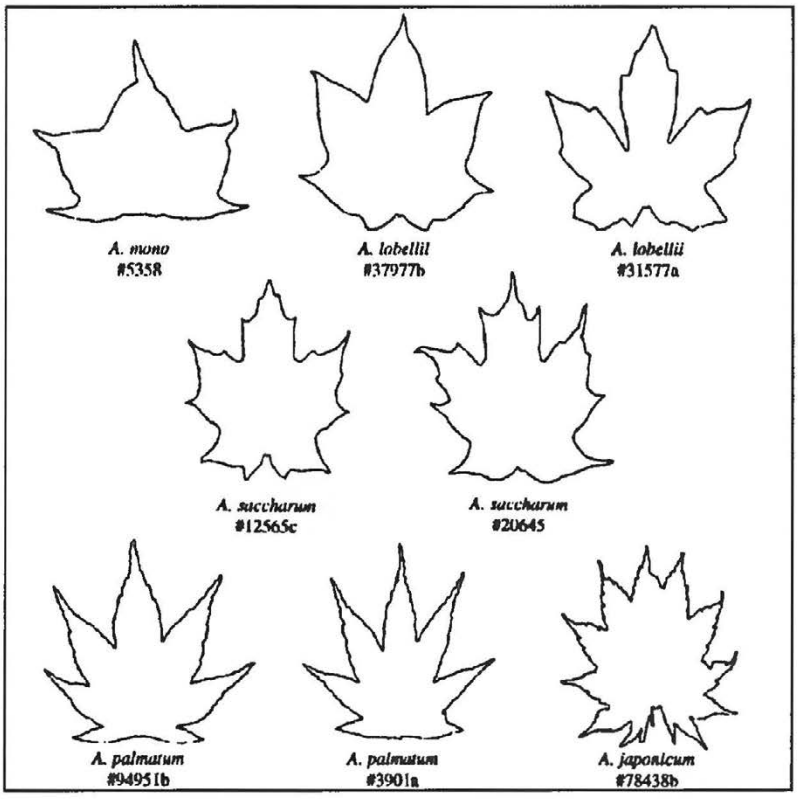


Fig. 2
Sample leaves from each of the eight trees of the genus Acer discussed in the text. Numbers are Arnold Arboretum identification numbers

Table 2

Sample means and standard deviations for D estimates based on step lengths between 0 and 403 "units" (see text for details). Sample size, N, includes five replicate measurements of each leaf. Actual number of individual leaves shown in parentheses

Species	Tree #	Mean D	Std. Dev.	N
Acer mono				
	5358a	1.025	0.013	240 (48)
A. lobellii				
	37977b	1.028	0.011	145 (29)
	31577a	1.036	0.016	235 (47)
A. saccharum				
	12565c	1.043	0.020	190 (38)
	20645	1.056	0.020	145 (29)
A. palmatum				
	3901a	1.077	0.015	90 (18)
	94951b	1.078	0.012	130 (26)
A. japonicum				
	78438b	1.123	0.017	95 (19)

Table 3

Analysis of variance table for D estimates based on step lengths between 0 and 403. Mean values are reported in Table 2

Source	df	SS	Fs	p
species	4	1.0244	35.22	0.0074
tree (species)	3	0.0218	6.81	0.0002
leaf (tree(species))	246	0.2625	15.99	0.0001
error	1016	0.0678		
total	1269	1.3534		

How the variability in D values is distributed across various levels of the analysis is summarized in Figure 3. This shows the amount of relative variability attributable to differences in trees within a species, leaves within a tree, and that due to measurement error (the replicated measurements of D from random starting points). The species level is not included since it is considered a fixed treatment effect and differences at this level can, at least conceptually, be made arbitrarily large.

The pattern of the distribution of variability indicated by the variance components is quite plausible; just over twenty percent of the variability is attributable to measurement error, nearly sixty-five percent to variation of leaves

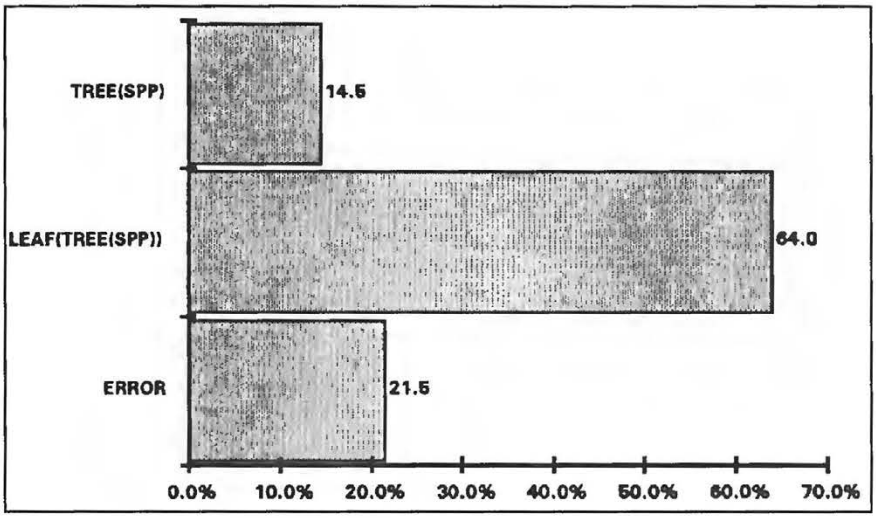


Fig. 3
Variance components for D estimates using all step lengths between 0 and 403

within a particular tree, and about fourteen percent to variation between trees within a species. Similarly, the analysis of variance is reassuring in that species with obviously different outline complexities (Fig. 2) have significant species-level differences in the ANOVA. Higher D values are associated with more complex outlines, and measurement error is sufficiently small so as not to obscure differences between individual leaves.

Despite their apparent plausibility, a closer examination of the results reveals several methodological inadequacies that could invalidate the results - hence the heading "naive analysis." Experience has shown that although these problems, plus a few more, are obvious once revealed, they are sufficiently subtle as to be commonly overlooked. Furthermore, they are sufficiently pernicious as not to easily admit a universal solution.

The main problems with the analysis are revealed in Figure 4. This figure shows a plot of log mean length estimates versus log stepsize for each tree. (Actual D estimates were made using individual, not mean, length estimates. Means are shown here for simplicity.) Differences between species are readily apparent with the highly complex *Acer palmatum* and *A. japonicum* forming a band across the top of the plot and the simpler species occupying the lower portion.

Recall that in the discussion of fractals and the development of the technique to estimate D, the key feature was that as one increased the resolution, i.e. decreased step-size, the number of steps required to traverse the contour increased (Equation 1). This results in increased length measurements using smaller length scales. Notice, however, in Figure 4 that for log step lengths smaller than about 2.0 the log of the length estimates (and the original estimates) manifest a more or less consistent decrease.

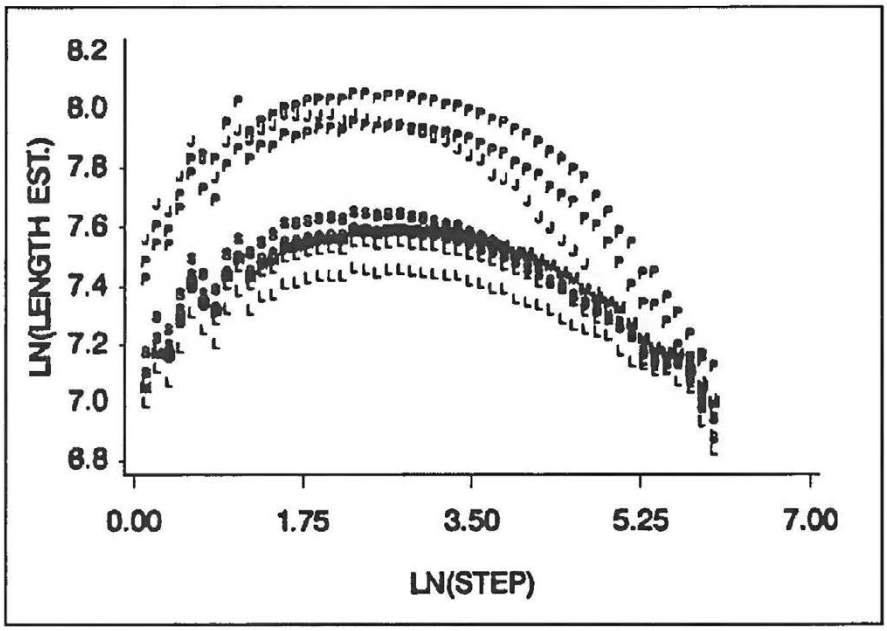


Fig. 4
Average log length estimate versus log step size for each tree. Species are distinguished by symbols: J = Acer japonicum, L = A. lobellii, M = A. mono, P = A. palmatum, S = A. saccharum

The decrease in length estimates with decreasing step lengths in the figure is the result of inappropriately chosen step sizes. One must limit the analysis to the use of step lengths that are not only within the range of biological interest but also do not exceed the resolving power of the available data. In this case, the smallest step length for which one could obtain a meaningful length estimate was originally limited by the video image to one pixel. However, each outline was enlarged to a standard area to limit the effects of size differences on the analysis. This increased the minimum resolution to about 1.5 pixels. That magnification was not considered in the construction of the original set of step lengths, and the pathological parts of the curves in Figure 4 are the result of using step lengths smaller than the minimum resolution of the data. In such cases, every adjacent point in the outline exceeds the step length and will increment the step count. The resulting length estimate is simply the number of points in the outline multiplied by the length of the step. Smaller steps give smaller lengths. Differing outline complexities still give significantly different D, but the artifactual nature of the curve at small scales requires they be omitted to obtain better estimates of the fractal dimension. In fact, as will be discussed later, a minimum step size of ten times the minimum resolution of the data is desirable. That would suggest a minimum step size of 15.0 units (2.71 log units) in the present case.

The importance of inappropriately small step lengths exceeds the relative number of such values used in the analysis. This is because the estimate of D is based on a linear regression and is thus especially sensitive to the influence of values of the independent variable which are far from their mean. If the effects were distributed randomly across the range of step lengths then D might be relatively unaffected. However, the very nature of the problem concentrates the erroneous length estimates on the left-hand side of the trajectory and thus magnifies their effect on the estimated slope.

Similar problems can be associated with large step sizes. These and other problems illustrated by Figure 4 will be discussed shortly. For now, we can redo the analysis using only step lengths between 15.0 and 200 (2.71-5.30 log units) to see what effect that has on the results.

Reanalysis

Mean values of D for each species based on the new analysis are given in Table 4. In each case, there is an increase in estimated mean D over the earlier results, but the ranking of trees by D remains unchanged except for the *Acer lobelii*. The elimination of the too-small step lengths has allowed the regression to fit the curve more closely and produce better estimates of D. The values for *Acer saccharum*, 1.163 and 1.188 are similar to the value of 1.18 found by Vlcek & Cheung (1987) for the same species.

Table 4
Sample means and standard deviations for D esimates based on step lengths between 15 and 200

Species	Tree #	Mean D	Std. Dev.	N
Acer mono				
	5358a	1.140	0.027	240 (48)
A. lobellii				
	37977b	1.135	0.022	145 (29)
	31577a	1.142	0.027	235 (47)
A. saccharum				
	12565c	1.163	0.032	190 (38)
	20645	1.188	0.040	145 (29)
A. palmatum				
	3901a	1.218	0.034	90 (18)
	94951b	1.221	0.019	130 (26)
A. japonicum				
	78438b	1.309	0.029	95 (19)

The use of the restricted step sizes has relatively little effect on the significance levels in the analysis of variance (Table 5). Both effects of tree-within-species and leaf-within-tree are still highly significant $p<0.001$, although the tree-within-species probability increased from 0.0002 to 0.0008. The significance of the species effect is increased with its probability level decreasing to 0.0061 from its previous value of 0.0074.

Table 5

Analysis of variance table for D estimates based on step lengths between 20 and 200. Mean values are reported in Table 4

Source	df	SS	Fs	p
species	4	2.9048	40.28	0.0061
tree (species)	3	0.0541	5.79	0.0008
leaf (tree(species))	246	0.7665	10.77	0.0001
error	1016	0.2940		
total	1269			

The nature of the differences between this and the earlier analysis is most revealed in the estimation of the variance components (Fig. 5). Most of the variability in the random effects is still associated with leaves within trees, but Figure 5 indicates variability attributable to error has increased by 50%. In this case, error is the difference in the estimation of D for individual leaves based on random starting locations. This is understandable since repeated measures of the same leaf with step lengths below the minimum resolution of the data must yield identical length estimates. These estimates would act to pull the regression line down on the left toward an essentially constant value and thus mask the variability due to starting location found in the other parts of the curve.

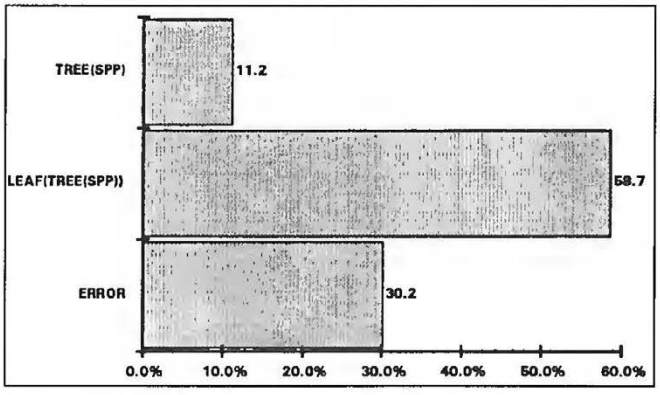


Fig. 5

Variance components of D estimates using step lengths beween 15 and 200

Figure 4 revisited

There are a few more points that can be made from Figure 4. First, one notices that length estimates for all species appear to converge to more or less similar values for larger step lengths. As step length increases fewer steps are required to traverse the outline until only one or a few steps can be made. Since the leaf outlines were scaled to a common area, the upper limit for step lengths that require a nonzero number of steps is about the same for all leaves. A similar number of large steps for all outlines yields similar length measurements. In fact, at a certain scale the leaf outlines are indistinguishable from circles. Unlike the convergence of length estimates due to inappropriately small step lengths, this convergence is driven by the actual shape of the leaves. Still, one may not wish to confound the general roundness of all the data with more interesting aspects. The use of 200 pixels as a reasonable upper limit was based on considerations discussed in the next section.

Another important consideration in fractal analysis is found between the minimum resolution suggested for this data and the upper limit where lengths converge. The path taken by each species differs not only in slope but in curvature as well. Length estimates for *Acer japonicum,* for instance, are nearly identical to those for one of the *A. palmatum* trees in the area of a log step length of 2.7, but its length estimates fall off in a more linear manner. The *A. palmatum* trees follow a more curving trajectory.

These differences in trajectories of log length estimate through the log step lengths indicate that the assumed linear relationship within leaves does not apply. A closer examination of the trajectories reveals that the deviation is not haphazard and is itself open to interpretation. In species that are represented by more than one tree, i.e. *Acer palmatum, A. saccharum,* and *A. lobelii,* the conspecific trees have trajectories that appear more similar to each other than to those of other species. For all species, the greatest agreement in trajectory is found in the area of smaller step sizes. The trajectories are rather flat and species are distinguished by their relative position on the log length axis. Yet, at the larger step lengths, length estimates tend to converge. The differences in D between species are thus due to differences in trajectory from the relatively flat, small scale portion to the similar large scale results. The differences we perceive in the complexity of leaf outlines is therefore concentrated at intermediate scales where the degree and pattern of "toothing" or additional lobing is manifest.

SPECIAL CONSIDERATIONS

The previous discussion of the maple leaf analysis emphasized a few of the methodological problems that arise in a fractal analysis. I would now like to present an overview of these and other problems that are of importance to any study

using the fractal dimension. Unfortunately, there are no universal rules for addressing these problems. The most desirable way to proceed will generally be dictated by the nature of the data or by experience with preliminary analyses. Emphasis is placed on outline analysis using the divider method, but the questions can generally be extended to any fractal analysis technique.

Data resolution

The precision of the data used for a fractal analysis is of more importance than in most other types of analysis. This is because D involves the examination of data over a range of scales. A precision that would be adequate for simple measurements at one scale may be insufficient at smaller scales to allow stable estimates of the fractal dimension. Of course, one is limited by the nature of the data and the equipment available for digitizing, but every effort should be made to collect data with sufficient resolution to ensure no loss of information within the range of scales of interest.

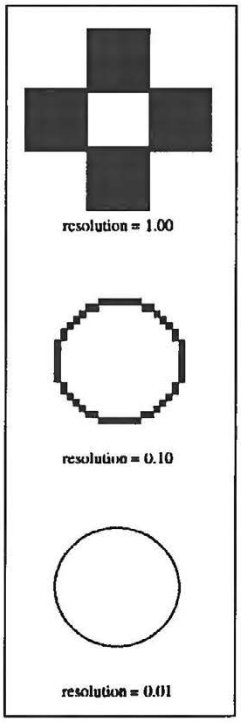


Fig. 6

A unit circle digitized using resolutions of 1 (top), eq 1/10 (middle), and eq 1/100 (bottom). When the resolution of the digitizer is close to the "size" of the object the pattern of the digitizing grid can impose artificial structure on the data

Figure 6 shows the results of the simulated digitizing of a circle at several resolutions. The form of the data at the lowest resolution (Fig. 6, top), where the resolution is on the same scale as the radius of the circle, obviously does not reflect the smooth shape on the original circle. It is instead dominated by the structure of the underlying digitizing grid. Increasing the resolution of the digitizer by a factor of ten still does not produce results that would be judged adequate, at least by eye (Fig. 6, middle). Only when digitizing resolution is increased to one one-hundredth the radius of the circle does its actual structure not appear dominated by that of the grid. This comparison suggests at an intuitive level a digitizing resolution of two orders of magnitude finer than the outline radius for closed, roughly circular contours. While this may seem a rather trivial exercise, it is easy to lose sight of these types of problems once the data have been neatly stored in a computer file.

Step size distribution

Another fundamental question that must be asked is what step sizes should be used? This is actually three questions in one: over what range will the step lengths be distributed, how many steps lengths should be used, and how should that number of steps be distributed over that range? Unfortunately, there is no simple answer to any of these.

The range of step lengths to use is a function of the resolution of the digitized data and the range of scales of interest to the researcher. The latter is not easily quantified. One may be interested in characterizing leaf shape from the scale of the whole leaf down to that of the smallest serration or from the level of a single marginal tooth down to that of the cellular structure of the margin. While both of these ranges are intuitively obvious, they are a bit more difficult to quantify, especially for complex forms. The lower limit at which the data can be analyzed will usually be set by the digitizing process. In any case, one would like to determine the range of step lengths in a manner independent of a particular outline or data set so that meaningful comparisons can be made to results obtained from other studies and data sets.

One approach to more data-independent determination of step range is to use a generalized model and base step-range selection on the model instead of the actual data. In the case of leaves or other closed outlines, a reasonable model is a circle. A data range that gives appropriate results for the circle, D=1, should then give results for the data that describe their how their complexity differs from that of a circle with a fair degree of accuracy.

The circle has a convenient property that especially lends itself to being a model for the fractal analysis of closed outlines. The exact number of steps, $N(\lambda)$, of length λ required to traverse a circle can be obtained from the formula for the side length of an n-sided polygon inscribed within the circle:

$$\lambda = 2R \sin(\pi/n),$$

where R, is the radius of the circle and n the number of sides. From this equation one can determine the formula for the fractal curve that would be obtained using the divider method:

This curve is plotted for a range of step lengths (0,2R] in Figure 7 and shows

$$\hat{L} = N(\lambda)\,\lambda = \frac{\pi}{\sin^{-1}(\lambda/2R)}\,\lambda.$$

that the assumed linear relationship between log length estimate and log step length of Equation 3 is badly violated for step lengths approaching the diameter of the circle. The curve is approximatly linear only at relatively small step lengths.

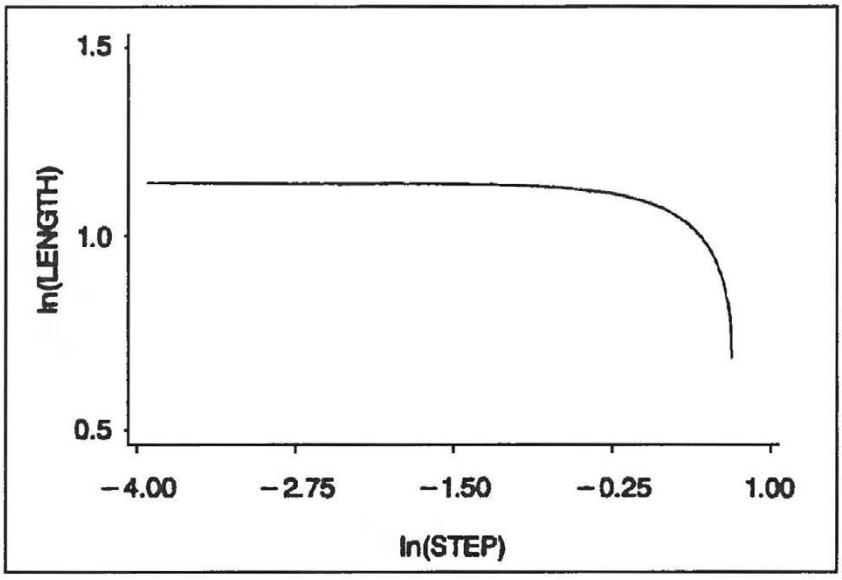


Fig. 7

Fractal estimation curve for a unit circle. The linearity assumed in the estimation procedures is badly violated for step lengths approaching the diamter of the circle

One approach to measuring the fractal dimension of a circle would be to use only step lengths up to the length of its radius. Extending this to the outlines of real objects and assuming they have been scaled to common area, one could

set as the upper limit for step length the radius of a circle of the same area, $\lambda_{max} = \sqrt{A/\pi}$, where A is the area. This length would require six steps to cover the circle and at least that many for more complicated outlines since a circle has the shortest perimeter for a given area. This is the basis for choosing a maximum step length of around two hundred pixels used in the reanalysis of our leaf data which eliminates a fair amount of the region at larger step sizes in Figure 4 where the length estimates converge. For nonstandardized data one might use the radius of a circle with the same area as the smallest outline.

The smallest step sizes that can be used will often be limited by the resolution of the digitizer. While this would seem to suggest that the minimum resolution of the digitizer would be an appropriate limit for the smallest step length, simulations show that orientation on the digitizing grid can effect length estimates at small scales (Slice, unpublished). The length estimates of straight lines digitized through a series of angles with respect to the digitizing grid show effects of orientation that are a function of the cosine of the angle. These effects are negligible for step lengths ten times the scale of the digitizing resolution, and I have adopted that as the minimum step length for my own work. Using this "rule of thumb", the minimum step length for data scaled to a common area would be ten times the resolution of the digitizer multiplied by the largest scale factor used on the data.

Applying these criteria to a digitized circle results in average D estimates of 0.995 with a range of 0.990 to 0.999 - very close to the theoretical value of 1.

If the data under consideration are perfect fractals, one need only use two step lengths to determine D. This is generally not the case. The inherent variability in most data requires multiple step lengths be used to ensure good estimates. The exact number to use is not readily apparent. If the analysis is carried out by hand by actually stepping around a drawing or map with a divider, then the laboriousness of the process will surely limit the number of step lengths that can reasonably be used. For computer analysis, this is less of a problem and the required number of steps is a question of how thoroughly one wishes to cover the step range. Too few steps could prevent the detection of interesting nonlinearities and too many simply consume computer time while not producing additional useful information. The exact number is probably best determined by examining plots of the length-step relationships of subsamples of the data. The fifty step lengths used in the leaf analysis were able to describe the important features of the fractal estimation curves and identify problems with the range of steps. The twenty-two lengths within the appropriate range appear to be sufficient to characterize the estimation curve for different leaves.

Finally, one must decide on the distribution of the step lengths across their range. The simplest decision would be to evenly space the steps between the minimum and maximum values. The estimation process, however, operates on log transformed results and a set of evenly spaced step lengths, when log transformed, will be highly concentrated in the lower values of the range. This

will tend to give more weight to the greatest lengths in determining the slope of the regression line. Alternatively, one can distribute the step lengths so that after log-transformation they are evenly distributed across the range of log step lengths.

Starting location

The point on the outline from which steps are counted can have an effect on the estimation of D. In some cases, an outline may have curves positioned in just such a way that they are included in the length estimate when starting from one point, but not when starting from another. Numerous situations like this are likely to be found in any reasonably complex curve, but are probably more important when they involve larger features since they mainly effect results for larger step sizes which in turn have considerable influence on the estimate of D.

The simplest way to address the effect of starting point on D is to analyze each curve using a number of different starting locations. The mean of the resulting estimates can then be used as the estimate for the fractal dimension of the contour and their variance and distribution used to assess error due to starting location. The importance of this type of effect is seen in Figure 5 where a great proportion of within-tree variability in D estimates (over 30%) is found within leaves.

The number of different starting points to use will depend largely on the data. The smooth, symmetric structure of a circle is such that little if any replication would be required. In the leaf analysis, five estimates were used for each leaf. This number was determined primarily by resource constraints. As a general practice, one could run a series of analyses on a sample data set using different numbers of random starting locations and plot the within-leaf sample variance versus number of replicates. One expects the variance estimate to vary rather widely at first then converge on some parametric value. The number of replicates to use could then be determined as the minimum number for which the within-leaf variance is deemed sufficiently stable.

Standardization

To standardize or not to standardize? That is an important question. The reason is that consideration of whether or not to standardize closed outlines to a common area brings up questions about the underlying generating process. Nonfractal outlines with similar "shapes" have a constant perimeter-area ratio (actually, one considers the ratio of the perimeter to the square root of area to achieve a unitless number). In the case of circles, this ratio is always $2/\sqrt{\pi}$. For squares, it is 4. The situation is different for fractals. The area of a plane enclosed by a fractal

outline is constant, and increasing the resolution of its measurement will result in convergence to a finite value (this implies the box-counting method). Recall, though, that as measurement resolution increases the length of a truly fractal perimeter will diverge to infinity. This relationship is described by an equation due to Mandelbrot (see Feder, 1988) as $\rho = \text{Perimeter}^{1/D}\text{Area}^{-1/2}$, where ρ is constant for similar shapes, and perimeter and area are determined at some fixed step size. This equation can be used to formulate a relationship between perimeter length and area that can be used to estimate the fractal dimension of outlines,

$$P(\delta) = C\delta^{(1-D)}\sqrt{A(\delta)^D}. \qquad (4)$$

Here C is a constant and $P(\delta)$ and $A(\delta)$ are perimeter and area determined at a sufficiently small step size, δ.

There are two important points concerning perimeter-area relationships. First, while previous methods used differences in length estimates at different step sizes to estimate the fractal dimension of a single outline, the use of the perimeter-area relationship examines a number of outlines of different sizes to estimate D. One must therefore have access to a sufficiently broad range of outline sizes to effectively use this technique, and they must be assumed to have similar fractal shapes. One would, of course, not want to standardize the data in this case (the standardization is actually built into Equation 4). The second point is that this approach deals more with using the outlines to study the generating process than the outlines themselves. This may be appropriate for using the shapes of rain clouds to study atmospheric processes (Lovejoy, 1982) but perhaps not for biologists interested in quantifying the complexity of a particular part of a plant or animal.

When the complexity of a structure is the focus of study, one will probably want to standardize the data. The inverse of the scale factors can be retained as "size" variables and related to differences in complexity. For example, small, very young leaves could have the same shape as larger specimens or they could become more (or less) complicated as they matured. If they were the same shape and small and large individuals differed only by an isometric scaling, standardization would produce similar estimates of D. If the linear relationship assumed in Equation 3 does not hold, then standardization becomes necessary. Otherwise, different parts or proportions of the nonlinear curve will appear in the range of the log step lengths used. This can result in quite different D estimates for shapes with the same structure.

Another probably frequent relationship between size and shape would be small structures differing from larger ones by simple affine transformations. The utility of standardization in these cases would be questionable.

Nonlinearities

The fundamental assumption in the estimation of the fractal dimension as outlined above is the linear relationship between log length estimate and the log of the scale used for measurement. Quite different log length-step relationships can produce a given D value; the curve may be linear, it may have a lower slope for smaller step lengths and a compensatory higher slope at larger scales, or vice versa. This makes it highly desirable to examine the log-log plots to assess the importance of any nonlinearities. If any are present one may wish to perform additional analyses that would isolate linear regions of the curve and treat them separately. Alternatively, the trajectories may not be sufficiently similar across a sample to allow for such a partitioning. In this case, one would have to focus on interpreting the differences in curves for different specimens or groups.

This kind of effect was seen in the maple leaf data. The leaves generally have curves that were quite variable between groups. Some species showed rather flat overall relationships while others were more sloped at intermediate scales. All tended to converge to similar length estimates for higher values, but at different rates depending upon the initial location and structure of the curve. Other data could show consistent slope patterns for one range of step sizes, but differ at another. While there is nothing one can do to remove any nonlinearities (they are a property of the data itself), one must be aware of their presence and allow for their identification and possible interpretation.

CONCLUSION

The calculation of the fractal dimension is not as simple as weighing or measuring the length of an object, and being a summary measure, D can mask underlying patterns and relationships that may be important. Advances in data collection and computing methods will make the former increasingly less of a problem. As to the latter, neither do lengths have information about the relative contribution of constituent parts to the length of the whole, and weights tell nothing of the distribution of mass. In this respect they are no more adequate than D. Even in the early stages of its development, the application of fractal analysis to the study complexity in biological shapes has contributed to our understanding and quantification of things as diverse as the distribution of eagle nest sites, the morphology and growth patterns of whole organisms, and the shapes of parts of organisms such as the leaves discussed here. While D is not as frequently used a measure as grams or meters, the ubiquity of complexity in biological structures and its potential importance suggest it may become an important part of the language of biological shapes.

ACKNOWLEDGEMENTS

Early versions of the manuscript benefited from comments by F. James Rohlf, Jessica Gurevitch, and Scott Ferson.

The work was supported in part by the graduate program in Ecology and Evolution at the State University of New York at Stony Brook, NSF grant BSR8918630 to F. James Rohlf, NSF dissertation improvement award BSR9015837 to D. Slice, a Katherine Putnam Fellowship at the Arnold Arboretum of Harvard University to J. Gurevitch, and the 1991 Morphometrics Workshop, Valsain, Spain.

This paper is contribution number 824 from Graduate Studies in Ecology and Evolution, State University of New York at Stony Brook.

REFERENCES

BARNSLEY, M., 1988. *Fractals Everywhere*. Academic Press, Boston.

BRADBURY, R. H., R. E. REICHELT & D. G. GREEN, 1984. Fractals in ecology: methods and interpretation. *Marine Ecology*. Progress Series, 14: 295-296.

BURLANDO, B., 1990. The fractal dimension of taxonomic systems. *Journal of Theoretical Biology*, 146: 99-114.

CADDY, J. F. & C. STAMATOPOULOS, 1990. Mapping growth and mortality rates of crevice-dwelling organisms onto a perforated surface: the relevance of "cover" to the carrying capacity of natural and artificial habitats. *Estuarine, Coastal and Shelf Science*, 31: 87-106.

CALDWELL, C. B., S. J. STAPLETON, D. W. HOLDSWORTH, R. A. JONG, W. J. WEISER, G. COOKE & M. J. YAFFE, 1990. Characterisation of mammographic parenchymal pattern by fractal dimension. *Physics in Medicine and Biology*, 35: 235-247.

DICK, M. & P. A. BURROUGH, 1988. Using fractal dimensions for characterizing tortuosity of animal trails. Physiological Entomology 13: 393-398.

FEDER, J., 1988. *Fractals*. Plenum Press, New York.

FRONTIER, S., 1987. Applications of fractal theory in ecology. In *Developments in Numerical Ecology*. P. and L. Legendre, Eds. Springer-Verlag, Berlin.

GLENNY, R. W., H. T. ROBERTSON, S. YAMASHIRO & J. B. BASSINGTHWAIGHTE, 1991. *Applications of fractal analysis to physiology*. Journal of Applied Physiology, 70(6): 2351-2367.

HORSFIELD, K., 1990. Diameters, generations, and orders of branches in the bronchial tree. Journal of Applied Physiology, 68(2): 457-461.

KENT, C. & J. WONG, 1982. An index of littoral zone complexity and its measurement. *Canadian Journal of Fisheries and Aquatic Sciences*, 39: 847-853.

LOEHLE, C., 1983. The fractal dimension in ecology. Speculations in Science and Technology, 6(2): 131-142.

LONG, C.A., 1985. Intricate sutures as fractal curves. *Journal of Morphology*, 185: 285-295.

LOVEJOY, S., 1982. Area-perimeter relation for rain and cloud areas. *Science,* 216(9): 185-187.

LYNCH, J. A., D. J. HAWKES & J. C. BUCKLAND-WRIGHT, 1991. Analysis of texture in microradiographs of osteoarthritic knees using fractal signatures. *Physics in Medicine and Biology,* 36(6): 709-722.

MANDELBROT, B.B., 1983. *The Fractal Geometry of Nature,* 2nd. edition. W. H. Freeman and Company, New York.

MATSUYAMA, T., M. SOGAWA & Y. NAKAGAWA, 1989. Fractal spreading growth of *Serratia marcescens* which produces surface active exolipids. *FEMS Microbiolology Letters,* 61(3): 243-246.

MORSE, D. R., J. H. LAWTON, M. M. DODSON & M. H. WILLIAMSON, 1985. Fractal dimension of vegetation and the distribution of arthropod body lengths. *Nature,* 314: 731-733.

OBERT, M., P. PFEIFER & M. SERNETZ, 1990. Microbial growth patterns described by fractal geometry. *Journal of Bacteriology,* 172(3): 1180-1185.

PALMER, M.W., 1988. Fractal geometry: a tool for describing spatial patterns of plant communities. *Vegetatio* 75: 91-102.

PEITGEN H. & D. SAUPE, 1988. *The Science of Fractal Images.* Springer-Verlag, New York.

PENNYCUICK, C. J. & N. C. KLINE, 1986. Units of measurement for fractal extent, applied to the coastal distribution of bald eagle nests in the Aleutian Islands, Alaska. *Oecologia,* 68: 254-258.

PHILLIPS, J.D., 1985. Measuring complexity of environmental gradients. *Vegetatio,* 64: 95-102.

RITZ, K. & J. CRAWFORD, 1990. Quantification of the fractal nature of colonies of *Trichoderma viride. Mycological Research,* 94(8): 1138-1152.

ROHLF, F. J. & D. E. SLICE, 1990. Image: image enhancement program. Distributed as part of *Proceedings of the Michigan Morphometrics Workshop software.* F. J. Rohlf and F. L. Bookstein, Eds. Special Publication Number 2 University of Michigan Museum of Zoology.

SAS®, 1985. Statistical Analysis Software, version 5. SAS Institute Inc., Box 8000, Cary, North Carolina, USA 27511-8000.

SLICE, D.E., 1989. *Fractal-D.* Exeter Publishing, Ltd., 100 North Country Rd., Setauket, New York, USA.

TATSUMI, J., A. YAMAUCHI & Y. KONO, 1989. Fractal analysis of plant root systems. *Annals of Botany,* 64: 499-503.

VLCEK, J. & E. CHEUNG, 1986. Fractal analysis of leaf shapes. *Canadian Journal of Forest Research,* 16:124-127.

VOSS, R.F., 1988. Fractals in nature: from characterization to simulation. In *The Science of Fractal Images.* H. Peitgen and D. Saupe, Eds. Springer-Verlag, New York.

WEST, B. J. & A. L. GOLDBERGER, 1987. Physiology in fractal dimensions. *American Scientist,* 75: 354-365.

PART FOUR

APPLICATIONS

ONTOGENETIC ALLOMETRY OF THREESPINE STICKLEBACK BODY FORM USING LANDMARK-BASED MORPHOMETRICS

JEFFREY A. WALKER

Department of Anatomical Sciences
Health Sciences Center
SUNY at Stony Brook
Stony Brook, NY 11794
JWALKER@SBBIOVM. bitnet

CONTENTS

ABSTRACT

Qualitative studies of ontogenetic allometry are useful both for examining functional and ecological consequences of shape change during growth and for inferring historical processes that potentially explain the diversity in adult form among closely related organisms. Allometry of body form from a single sample of threespine stickleback from Cook Inlet Alaska is studied here using landmark data. Landmark-based morphometrics currently provide the most powerful techniques for studying shape. Digitized landmark configurations for each specimen were superimposed using the generalized resistant fit algorithm (Rohlf and Slice 1990). Variation in size and shape among the superimposed specimens was used to calculate trajectories of allometric shape change termed allometry paths. Superimposition of the allometry paths on the corresponding landmarks provides a simple, efficient, and intuitive approach for studying qualitative patterns of allometry. The method is complimentary to other landmark based methods that have studied allometry by superimposing outlines of mean shape forms from a series of discrete size classes. Interpretation of the allometry paths are compared between the set of specimens fit by the generalized resistant fit and the set of specimens fit by the two-point registration. It is suggested that if the generalized resistant fit is able to produce a reasonably "correct" superimposition, allometry paths calculated from the residuals from the resistant fit are more easily interpretable, in terms of ecological and functional implications, than the corresponding allometry paths from a two-point registration.

INTRODUCTION

Change in body form during growth (ontogenetic allometry) is an important property of all organisms because of the many biomechanical, physiological, behavioral and ecological variables associated with body size. This study examines a new graphical method for studying allometry, one that superimposes estimated paths of shape as a function of size change onto a set of landmarks that describe body form. The method described here, an extension of the graphical based analyses of shape prominent in landmark-based morphometrics, is a simple, efficient, and intuitive approach for interpreting qualitative patterns of allometry in organisms. The method is applied to a cross-sectional ontogenetic series of the threespine stickleback, *Gasterosteus aculeatus*.

Threespine stickleback are widely distributed throughout much of the lowland holarctic (Bell 1984). Variation in body form is common and much of this variation is microgeographic in scale. Body form may be very different in adjacent lakes or streams (Lavin & McPhail, 1985; Reimchen *et al.*, 1985, Francis *et al.*, 1986), within the same stream (Baumgartner, 1986, 1992), or within the same lake (Larson, 1976; McPhail, 1984, 1992; Baumgartner *et al.*, 1988). Threespine stickleback have three general life-history forms, a marine form inhabiting open ocean or estuaries year round, an anadromous form that enters freshwater streams only to breed and a resident fresh-water form. Allozyme evidence (Withler & McPhail, 1985) is consistent with the hypothesis, based on extensive zoogeographic evidence (e.g. McPhail & Lindsey, 1970; Bell, 1976, 1984), that much of the phenotypic variation among freshwater stickleback along the pacific coast of Canada and Alaska has been independently derived due to repeated colonization from an anadromous ancestor following the last glaciation. Much of the variation among populations can be considered phylogenetically and statistically independent, an important property for comparative studies (Harvey & Mace, 1982; Pagel & Harvey, 1988; Harvey & Pagel, 1991; Ridley, 1983; Felsenstein, 1985, 1988). Diversification of threespine stickleback into numerous body forms likely reflects trophic and locomotor adaptations to local environments (Lavin & McPhail, 1985, 1986, 1987; Taylor & McPhail, 1986). Comparisons of patterns of ontogenetic allometry among populations might be useful for studying the microevolutionary dynamics of small phenotypic radiations. The analysis in this paper is a preliminary study of

ontogenetic shape change a in radiation of threespine stickleback from Cook Inlet, Alaska.

One of the appealing aspects of many of the landmark-based methods of shape analysis is the graphical approach for exploring shape change, perhaps best exemplified by the thin-plate spline (Bookstein, 1989, 1991). Simple superimposition of specimens is a first step in exploring shape variation within and among samples (Rohlf & Bookstein, 1990). Relative warps (or an analogous principal components analysis of the 2px2p covariance matrix of Procrustes residuals) is an elegant method to explore factors of within population covariation among the landmarks (Bookstein, 1991; Rohlf, this volume). Alternatively, one can explore the relationships of landmark variation with external variables such as geologic time (Bookstein & Reyment, 1989) or ontogenetic time (Bookstein, 1991). The covariation between landmark variation and size, or allometry, can be approached by relative warps, where presumably, if general size is a factor, one of the relative warps will model this (Bookstein, 1991) or by explicitly treating size as a covariate. There are many goals in studies of allometry. Often, univariate studies of allometry seek to test a priori quantitative hypotheses of the relationship between two variables. Multivariate studies of allometry are useful as qualitative descriptions of size correlated shape change, which may then be used to form hypotheses that relate to the biology of the organism. Traditional and recently developed methods in multivariate allometry, including the methods of Jolicouer (1963), Mosiman & James (1979) and Darroch & Mosiman (1985), involve the interpretation of a table of allometric coefficients, component scores, or ratios. With a large number of variables, visualizing shape change of the whole form measured (organ or organism) becomes tedious and increasingly difficult. Graphical, landmark-based methods that provide a visual description of allometric shape change allow a simpler and more intuitive approach for exploring allometry than offered by distance-based methods.

Allometry paths can be used on data superimposed by any Procrustes technique. In this paper, I compare the functional and ecological implications of allometry paths superimposed on specimens fit by the two-point registration and specimens fit by the generalized resistant fit. The generalized resistant fit is one of several Procrustes superimposition methods (Rohlf, 1990; Chapman, 1990; Goodall, 1991), all of which superimpose landmark configurations by rigid translation to a common location, rigid rotation to minimize some criterion of fit, and scaling the configurations to a common size. The brief introduction to superimposition methods that follows describes the methods in two-dimensions (X and Y coordinates) although it is easily extrapolated into higher dimensions. Three general types of Procrustes superimposition are the two-point registration, or edge superimposition (Goodall, 1991), optimal (least squares) superimposition, and resistant (repeated medians) superimposition. The two-point registration fits

specimens by matching a single, homologous edge, or baseline, among the set of configurations. Specimens are translated to center the baselines, rotated to align the baselines horizontally and scaled to make each baseline unit length (Bookstein, 1986, 1987, 1991; Bookstein *et al.*, 1985). The transformed coordinate values of the non-baseline landmarks are called shape coordinates (Bookstein *et al.*, 1985). In least squares superimposition, digitized specimens are superimposed by centering the configurations at the origin, scaling the configurations so that the sum of the squared distances between all landmarks and the origin equals unity, and rotating each configuration to minimize the sum of the squared distances between homologous landmarks. The configuration centroid is the mean X and Y coordinate value over all landmarks. The landmark centroid is the mean X and Y coordinate value at each landmark. The configuration of landmark centroids is the mean configuration. The residual difference between specimen coordinate values and landmark centroids are termed Procrustes residuals. Least squares superimposition has the undesirable property of distributing local shape differences among objects evenly across all landmarks (Siegel & Benson, 1982; Siegel, 1982; Benson, *et al.*, 1982; Olshan, *et al.*, 1982). The resistant fit effectively resists this global spread of local residual variation if most of the shape differences occur at fewer than half of the landmarks (Siegel & Benson, 1982). Resistant superimposition differs from least squares superimposition in the use of repeated medians to calculate translation, scaling and rotation parameters (see Siegel & Benson, 1982; Rohlf, 1990; Rohlf & Slice, 1990 for the detailed methodology). As in least squares superimposition, residual coordinates from the configuration mean are Procrustes residuals.

METHODS

Sample

The sample was collected during summer 1990 from Picnic Lake on the Kenai Peninsula near Cook Inlet, Alaska. The specimens were collected with minnow traps, baited with sharp Cheddar cheese and set overnight near vegetation within a few meters of shore. The trapped sticklebacks were anesthetized in the field with MS-222 (Tricaine Methanesulfonate) and fixed in 10% formalin. In the laboratory, the specimens were stained with Alizarin Red S and preserved in 50% isopropyl alcohol. From the large sample, 66 male specimens and 10 unsexed fry were chosen for the analysis to attempt to represent the size range of the sample. Sex was scored after measuring by making a small slit in the abdomen and inspecting the morphology of the gonad. Fry were not sexed because of the undifferentiated gonad at this small size.

Measurements

Fifteen landmarks on each fish were digitized (Fig. 1). These landmarks are: tip of upper jaw (premaxilla) (LM1), posterior border of the supraoccipital on the dorsal midline (LM2), anterior junction of the first dorsal spine with the dorsal midline (LM3), anterior junction of the second dorsal spine with the dorsal midline (LM4), junction of the first fin ray of the dorsal fin with the dorsal midline (LM5), insertion of the dorsal fin with the dorsal midline (LM6), origin of the caudal fin on the dorsal midline (LM7), caudal border of lateral line (LM8), origin of the caudal fin on the ventral midline (LM9), insertion of the anal fin on the ventral midline (LM10), junction of first fin ray of anal fin on the ventral midline (LM11), caudal tip of posterior process of pelvic girdle (LM12), posterior tip of ectocoracoid (LM13), anterior border of ectocoracoid on ventral midline (LM14), posterior edge of angular (LM15). These landmarks summarize the two-dimensional form of the lateral aspect of the fish. LM2, LM12, and LM14 cannot be located from a lateral view of the specimen and were pinned with small insects pins. The caudal border of the dorsal and anal fins were also pinned to aid in more precisely identifying these landmarks.

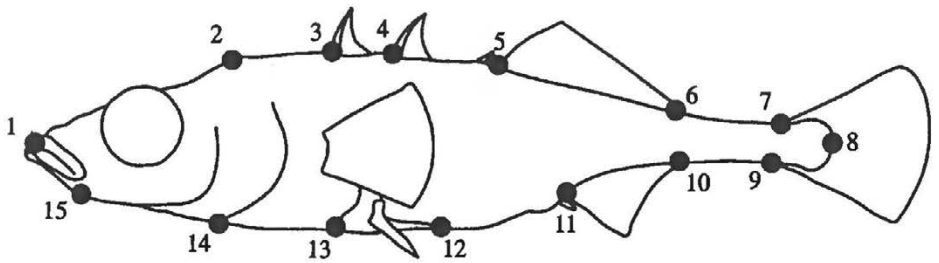


Fig. 1
Anatomical positions of the fifteen landmarks analyzed in this study

The landmarks were digitized using a CCD video camera with a 70mm macro lens and MorphoSys software on an IBM-compatible personal computer. The digitized coordinates were superimposed with the generalized resistant fit option of the Procrustes superimposition software, GRF (Rohlf & Slice, 1990). Superimposition methods have been criticized because of the number of different loss functions one may use to fit specimens (Lele, 1991). This critique is not unique to superimposition analyses but applies, in general, to all statistical methods. Different superimposition methods can indeed produce different fits (Chapman, 1990). This caveat should make one cautious about the technique used and, perhaps, explore differences among the loss functions, but does not justify rejecting these powerful techniques.

Allometry Paths

Several, generally complimentary, methods are available for describing allometry graphically. In general, allometry can be described by superimposition of outlines of the mean shapes of discrete size classes (e.g. MacLeod & Kitchell, 1990), age classes (e.g. Bookstein, 1991), or life-history stages (e.g. Reilly, 1990), or by allometric trajectories at each landmark. The outline method is useful if one wants to compare shape differences at specific ages or life-history stages. Trajectories are most useful if one wants a simple picture of allometry without any information about shape at specific stages.

Outlines or trajectories may be calculated from either a series of mean shapes from different size classes or from a regression that estimates mean form at a specific size. Previous studies of landmark-based allometry, using either outline or trajectory methods, have used only the mean form at discrete ontogenetic stages to determine the outlines or trajectories. Olshan *et al.* (1982), for example, studied longitudinal growth of the cranium of an individual macaque by superimposing the digitized landmarks of the skull at two different ages using the resistant fit algorithm developed by Siegel & Benson (1982). Residual vectors drawn between corresponding landmarks for the two ontogenetic stages represented the trajectories of growth related shape changes within the individual. Similarly, Reilly (1990) analyzed cranial shape differences among larval, metamorphic and post-metamorphic salamanders. Mean landmark configurations of multiple specimens within a single ontogenetic stage were compared between stages pairwise using resistant fit Procrustes superimposition. Reilly combined an outline of the ontogenetically earlier stage with a residual vectors to represent allometric shape differences. MacLeod & Kitchell (1990) superimposed outlines of mean shapes from eight size classes of Eocene foramanifera, although tests of linear allometry proved to be insignificant (see below). Finally, Bookstein (1991) examined calvarial growth in the rat by superimposing outlines of mean forms of eight age classes using the two-point registration.

Division of a continuous growth series into discrete stages is most useful if one is interested in comparing shapes of specific stages, as in Reilly (1990) or Bookstein (1991). If allometry is nonlinear, arbitrary division of a continuous series into discrete size classes can potentially lose useful allometric information. This problem is most acute if only two size classes are compared. Division of the size series into multiple size classes reduces this information loss and should converge on trajectories estimated from regression. This can easily be seen by comparing figure 2, the allometric trajectories calculated from quadratic regressions of shape coordinate on log centroid size, for the sample of 164 rat skulls (ignoring longitudinal information) with figure 7.6.7a of Bookstein 1991. The initial roll back and subsequent roll forward of the vault is readily

demonstrated with the allometric trajectories. It should be apparent that if mean shape for a size class is calculated from only a few specimens, outlines or trajectories calculated from these mean forms can potentially give misleading results because of poor estimation of means.

Fig. 2
Quadratic allometry paths fit through combined sample of 21 rats measured at eight growth stages. Compare to figure 7.6.7a of Bookstein (1991)

Trajectories and outlines describing allometric shape change can be calculated easily and efficiently by independently regressing the two Procrustes residuals at a landmark (or shape coordinates if one is superimposing via the two-point registration) on S, where S is centroid size, the sum of the squared distances from each landmark to the centroid of its configuration. Centroid size is a useful proxy for General Size (Bookstein *et al.*, 1985) because, on the assumption of equal and uncorrelated variation of residuals from landmark centroids, the appropriate null model is one in which Centroid Size is orthogonal to the space spanned by the residuals (Bookstein, 1986). I have termed a trajectory calculated from regression of Procrustes residuals on S an allometry path (AP). After estimating regression parameters of both X and Y Procrustes residuals, the AP for a landmark can be reconstructed by a series of points whose coordinate values, (X_i,Y_i), are the expected values at S over the range of S for the specimens in the analysis.

Variations

A potential caveat in reconstructing AP's results from the many regression models one might chose to estimate regression parameters. The thirty scatterplots of S (abscissa) and all thirty Procrustes residuals (ordinate) from the fifteen landmarks in this study are illustrated in figure 3. Because the expectation of covariation

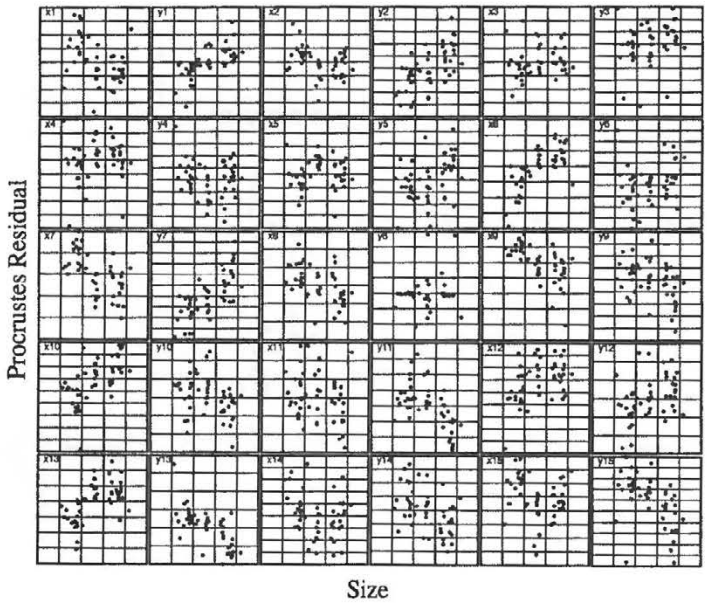


Fig. 3
Scatterplots of Procrustes residuals (ordinate) and Centroid Size (abcissa) for the fifteen landmarks

between Procrustes residuals and S is zero, isometry would be characterized by a scatter with no relationship between residuals and centroid size. Perusal of figure 3 indicates that allometry at these landmarks is common. Furthermore, shape appears to vary both linearly and nonlinearly with size. Some nonlinear allometries look asymptotic while others look quadratic. Reconstruction of AP's should take into account potential nonlinear allometries of landmarks. Several commonly used methods to estimate linear regressions are ordinary least squares, major axis and reduced major axis (Kuhry & Marcus, 1977; Harvey & Mace, 1982; Seim & Saether, 1983; Ricker, 1985; Rayner, 1985; McArdle, 1988; Riska, 1991). Ordinary least squares regression of the Procrustes residuals (and shape coordinates) on log transformed centroid size was used in this analysis as the independent variable, S, is an explicitly derived proxy for the latent General Size factor that explains allometries (Bookstein *et al.*, 1985; Crespi & Bookstein, 1989; Bookstein, 1991). Two alternative linear methods that are not employed in this study but are more in the spirit of resistant Procrustes superimposition are:

1) minimize $\sum_{i=1}^{n}(Y_i - blogS_i)$ where $b = \sum_{i=1}^{n} \frac{(logY_i - median\ (logY))\ ((logS_i - median\ (logS))}{(logS_i - medion\ (logS)^2)}$, and

2) robust regression using repeated medians (Siegel 1982). Non-linear allometries estimated from polynomial regressions have been explored but are not reported in this paper. Linear, quadratic and cubic regressions are implemented in a program written for the Macintosh.

RESULTS

The 95% confidence ellipses on the bivariate distribution of Procrustes residuals for each of the landmarks are illustrated in figure 4. The most variable landmarks are the tip of snout and angular, the landmarks of the ventral border (excepting the anterior tip of the ectocoracoid), and the landmarks of the caudal peduncle. The landmarks of the caudal region vary much greater in the anteroposterior direction than in the dorsoventral direction. Of particular interest, variation in the snout (LM1) is large and principally in a anterodorsal-posteroventral direction, while variation at the caudal tip of the lateral line (LM8) is nearly all in the anteroposterior direction. Figure 3 shows the scatterplots of all thirty Procrustes residuals against centroid size. Using standard F-tests (Sokal & Rohlf, 1981), a least squares regression of the Procrustes residuals on log S is significant ($a = 0.05$) for 21 of the 30 coordinates and significant linear allometry occurs in 13 of the 15 landmarks. Only LM3 (first dorsal spine) and LM4 (second dorsal spine) do not show significant allometry in the size range examined here. Perusal of the scatterplots indicate that shape change at many landmarks is nonlinear. A higher order (quadratic) regression is significant for 24 of the 30 Procrustes residuals and explains significantly more variation than

a linear fit in 10 Procrustes residuals, six of which have significant linear fits. Significant allometry, either linear or quadratic, occurs in all landmarks except LM3. The strongest allometries occur at the landmarks of the head, median fins, caudal peduncle, and ventral border.

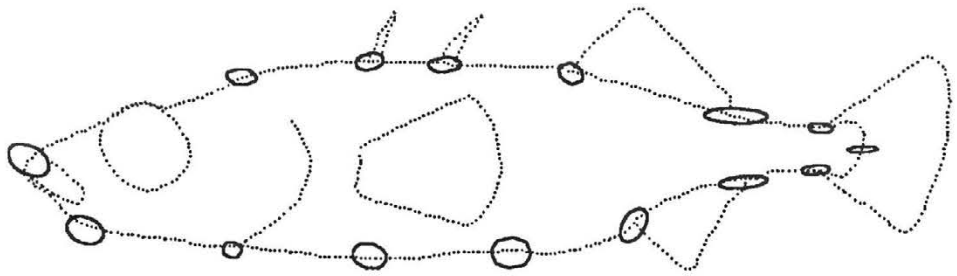


Fig. 4
Bivariate confidence ellipses of residual variation of fifteen landmarks for 76 specimens superimposed by generalized resistant fit

Allometry paths calculated from the Procrustes residuals are illustrated in figures 5. The strength of allometry is indicated by the length of the AP. The magnitude of the AP includes only the line segment; the arrow indicates only the direction of shape change from small specimens to large specimens. It should be emphasized that errors in the estimation of the regression parameters effect both the length and the direction of the AP. Figure 5 suggests a number of significant trends in ontogenetic allometry of the picnic lake stickleback. First, the snout tends to lengthen both anteriorly and dorsally during ontogeny. Relative head size expands both dorsally

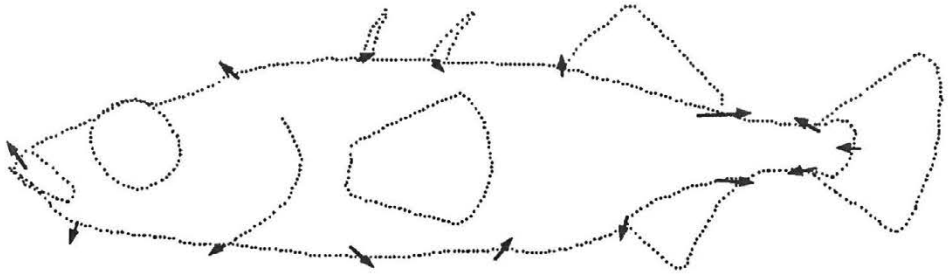


Fig. 5
Linear allometry paths at fifteen landmarks for 76 specimens superimposed by the generalized resistant fit

(greatly) and ventrally (moderately). Body depth increases ventrally (greatly) at the mid-body and both dorsally and ventrally at the median fins. The insertion of the dorsal and anal fins migrate caudally. Finally, the caudal peduncle both shortens in length and increases in depth (equally in the dorsal and ventral direction).

DISCUSSION

Perhaps the most interesting pattern of variation indicated in Figure 4 is the large dorsoventral variation in the snout, which is strongly correlated with size change (Fig. 3) and the extremely small dorsoventral variation at the caudal tip of the lateral line. The position of the mouth is highly variable among populations of threespine stickleback and the pattern of interlocality variation is similar to the intralocality variation indicated in this study (Walker in prep.). Variation in mouth position is well known in fish (Keast & Webb, 1966). In threespine stickleback, Hart & Gill (1992) have suggested that dorsoventral variation in mouth form may indicate foraging differences among populations: benthic and stream forms having more ventral mouths and limnetic forms having more dorsal mouths. The small dorsoventral variation at the posterior tip of the lateral line was generally expected as large variation in this direction might limit swimming performance. Additionally, the small dorsoventral variation at the posterior tip of the lateral line may indicate the general efficacy of the generalized resistant fit to superimpose these specimens into a biologically meaningful fit.

The patterns of allometry suggested by the allometry paths has interesting implications both for the ecology of the population and for interpretation of historical changes from a marine ancestor. In general, the head region expands during growth, which may improve trophic performance (acquiring and handling) of larger benthic prey by either increasing mouth gape and buccal volume or allowing for greater muscular development (Lavin & McPhail, 1985, 1986). Baumgartner *et al.* (1988) demonstrated patterns of interspecific shape variation in the head between a benthic-limnetic sympatric species pair of threespine stickleback very similar to allometry pattern in this sample. Comparisons of interpopulation differences in body form from multiple localities (Walker, in prep.) indicate that not only is interlocality variation similar to the pattern of allometry described by the allometry paths, but also that the anadromous form is very similar in head form to the smaller individuals of this sample. A possible exception to this increased benthic performance during growth is the anterodorsally directed allometry of the premaxilla.

Shape changes in the mid-body and caudal region are also in the direction of increased benthic design. It has been demonstrated theoretically, experimentally, and comparatively, that performance for rapid starts and turns, functions important for both escaping predators and foraging in complex habitats, is

augmented by deep bodies, caudally placed insertions of the median fins and short, deep caudal peduncles (see Webb 1982, 1984; Weihs, 1989 for review). These changes are precisely what is observed in the cross-sectional ontogeny of this sample of threespine stickleback. During growth in this population, the body deepens, the median fins migrate caudally, and the caudal peduncle both shortens and deepens. Inspection of individual landmarks reveal that, whereas, in the caudal peduncle region, dorsal and ventral landmarks contribute equally to the relatively shorter, deeper peduncle, the increased depth of the mid-body occurs largely from stronger allometry of the ventral landmarks. A study of the comparative locomotor performance between resident freshwater and anadromous threespine sticklebacks (Taylor & McPhail, 1986) demonstrated significantly greater fast start performance, deeper bodies and deeper caudal peduncles in freshwater specimens. Whether this pattern of allometry reflects natural selection of growth parameters or a phenotypically plastic response to environmental stimuli remains untested.

Concerning the method of allometry paths itself, these trajectories provide a method of graphically displaying and exploring biologic implications of allometric shape changes in body form that is complimentary to the outline method successfully used by Rielly (1990), MacLeod & Kitchell (1990) and Bookstein (1991). An outline method is more informative if the comparison of specific stages in the life-history of the organism are desired. Outlines may be misleading, however, if few size classes are compared and allometry is nonlinear or if the mean form of the size classes are not well estimated due to small sample sizes. In addition, superimposed outlines from numerous size classes may become cluttered with intersecting lines. Finally, superimposed outlines do not indicate the direction of shape change, although this may be easily resolved by labeling the lines with the size class as in Bookstein (1991). Allometry paths are useful for graphically displaying continuous size correlated shape change that may be easily compared among populations. For example, the superimposition of group mean forms and AP's for each group, with each group represented by different colors, offers a simple, exploratory method for comparing both mean shape and allometric differences among populations or species.

The use of the generalized Procrustes (either least squares or resistant) methods allow an efficient and intuitive approach for describing general patterns of allometric shape change in an organism. The emphasis of this approach was to develop a method for easily interpreting the functional and ecological implications of allometric shape change. Two other methods that have been used to analyze local allometries are polar coordinates (Ehlinger, 1991) and the two-point registration (MacLeod & Kitchell, 1990; Bookstein, 1991). Both of these methods require registering the specimens along a homologous baseline. Bookstein (1986) has demonstrated that for small shape changes, the choice of baseline does not effect the results of statistical tests for differences in mean shape among groups in the

two-point registration. For both methods, however, the direction and magnitude of allometric shape change at a landmark can only be interpreted with respect to the baseline. If shape changes normal to the baseline at either of the two baseline landmarks are relatively large, interpretation of results are, at best, ambiguous. For the specimens analyzed in this study, the obvious choice of a baseline would be the chord between LM1 (tip of premaxilla) and LM2 (caudal tip of lateral line) as in Ehlinger (1991). The Procrustes residuals indicate significant dorsoventral variation at the tip of the premaxilla but relatively little dorsoventral variation at the caudal tip of the lateral line. Because of this apparently large dorsoventral variation at the snout, the 76 specimens were refit by the two-point registration using the tip of the premaxilla (LM1) and the tip of the lateral line (LM8) as baseline landmarks. AP's for the thirteen non-baseline landmarks were estimated to compare to AP's of Procrustes fit specimens.

Bivariate confidence ellipses for the landmarks fit by a two-point registration are shown in figure 6 and should be compared with the confidence ellipses for the Procrustes residuals (Fig. 4). As expected, variation for the non-baseline landmarks is greater in the shape coordinates than in the corresponding

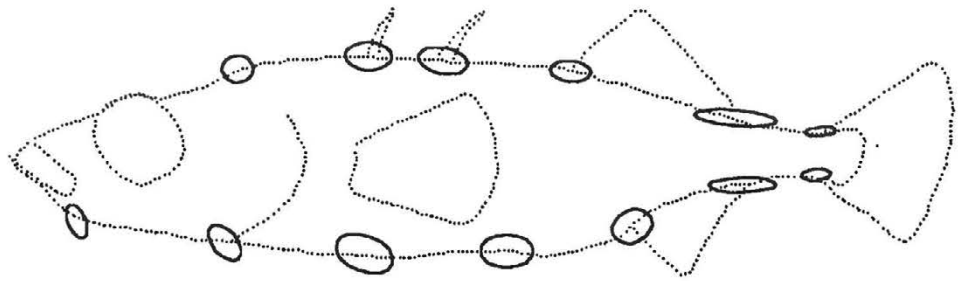


Fig. 6
Bivariate confidence ellipses of residual variation of thirteen non-baseline landmarks for 76 specimens superimposed by two-point registration

Procrustes residuals. By holding two landmarks constant, the two-point registration effectively transfers the variation at these landmarks to the other landmarks. Two-point registration allometry paths (TPRAP's) are illustrated in figure 7 and should be compared to the Procrustes allometry paths (PAP's) in figure 5. TPRAP's for the landmarks of the median fins and caudal peduncle were generally similar in magnitude and direction to the corresponding PAP's and are not discussed. Not surprisingly, there are predictable differences between the TPRAP's and the PAP's. The TPRAP for LM2, the caudal tip of the skull on the dorsal border is directed posteroventrally while the PAP for LM2 is larger in magnitude and directed anterodorsally. TPRAP's for LM3 and LM4, both on the dorsal border, have

moderate magnitudes and, as in LM1, are oriented posteroventrally. Linear allometries are not significant for the PAP's for LM3 and LM4. On the ventral border, TPRAP's at LM12 and LM13 have a similar orientation to the PAP's but are greater in magnitude. TPRAP's at LM14 and LM15, both oriented posteroventrally, are both greater in magnitude and differ in direction than the corresponding PAP's, which are oriented anteroventrally.

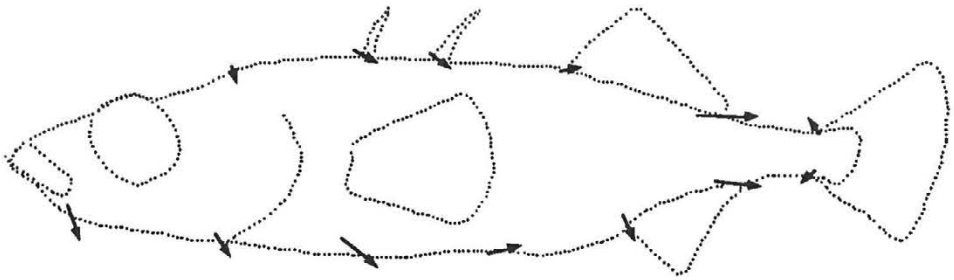


Fig. 7
Linear allometry paths at thirteen non-baseline landmarks for 76 specimens superimposed by two-point registration

Two general trends result from this comparison of TPRAP's and PAP's. First, with the exception of LM2, the magnitude of allometry tends to be greater in the TPRAP's. This greater magnitude must be due to the greater variation in the shape coordinates relative to the Procrustes residuals. Second, the more anterior TPRAP's all have a posterior component, while the corresponding PAP's have an anterior component. AP's of the landmarks fit by two-point registration must be interpreted in light of the constant baseline. The tendency for the orientation of all anterior TPRAP's to be directed posteroventrally at both dorsal and ventral landmarks may be real or due to the presence of allometry in the anterior baseline landmark (tip of premaxilla) oriented in the opposite direction (anterodorsal). If significant allometry exists in an anterodorsal direction at the tip of the premaxilla, the effect of removing all variation from this landmark will generally force the other landmarks to have posteroventral allometry (unless the real pattern of allometry is also dorsoventral, in which case, this real pattern will be canceled out by being forced in the opposite direction. This is probably the case in LM2). The pattern of allometry in landmarks fit by a two-point registration will necessarily be ambiguous, unless, one can be sure that the baseline is invariant among specimens. This is not to say that AP's of Procrustes fit specimens may not be misleading. The reality of PAP's is dependent on the efficacy of the generalized resistant fit in superimposing the specimens in a "correct" orientation, which cannot be known. But this is the case for the interpretation

of any scientific data; one never knows the truth. For many morphometric data sets, however, there is probably no a priori reason for suspecting that two landmarks are invariant. Although PCA of polar coordinates were not examined, it is expected that interpretation would be similar to that for the two-point registration. Whereas the two-point registration is reasonable for testing shape difference among specimens or groups, inference drawn from the direction of shape change is necessarily ambiguous because results have to be interpreted with respect to an invariant baseline. For this reason, fitting specimens with Procrustes superimposition offers a simple and efficient means to draw functional or ecological interpretations of size correlated shape changes.

ACKNOWLEDGMENTS

I'd like to thank M.A. Bell for providing both a lot of fish to study and the video and computer equipment needed to make the measurements. I thank M.A. Bell and Mark Spencer for discussing various aspects of this paper, F. Bookstein for critically evaluating previous versions of this manuscript, W.L. Jungers for the many discussions of size and shape, and F. James Rohlf for the many enlightening discussions of the recently developed morphometric tools.

REFERENCES

BAUMGARTNER, J. V., 1986. Unpublished Doctoral Dissertation. State University of New York at Stony Brook.

BAUMGARTNER, J. V., 1991. Spatial variation of morphology in a freshwater population of the threespine stickleback, *Gasterosteus aculeatus*. In Press.

BAUMGARTNER, J. V., M. A. BELL & P. H. WEINBERG, 1988. Body form differences between the Enos Lake species pair of threespine sticklebacks. *Canadian Journal of Zoology,* 66: 467-474.

BELL, M. A., 1976. Evolution of phenotypic diversity in Gasterosteus aculeatus superspecies on the Pacific coast of North America. *Systematic Zoology,* 25: 211-227.

BELL, M. A., 1984. Evolutionary phenetics and genetics. The threespine stickleback, *Gasterosteus aculeatus,* and related species. In B. J. Turner (eds.), *Evolutionary genetics of fishes.* Plenum, New York.

BENSON, R. H., R. E. CHAPMAN & A. F. SIEGEL, 1982. On the measurement of morphology and its change. *Paleobiology,* 8: 328-339.

BOOKSTEIN, F. L., 1986. Size and shape spaces for landmark data in two dimensions (with discussion). *Statistical Science,* 1: 181-242.

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.

BOOKSTEIN, F. L., 1991. *Morphometric Tools for Landmark Data. Geometry and Biology.* Cambridge University Press, Cambridge.

BOOKSTEIN, F., B. CHERNOFF, R. ELDER, J. HUMPHRIES, G. SMITH & R. STRAUSS, 1985. *Morphometrics in Evolutionary Biology.* Special Publication 15. Academy of Natural Sciences, Philadelphia.

BOOKSTEIN, F. L. & P. D. SAMPSON, 1987. Statistical models for geometric components of shape change. In A. S. Assoc. (eds.), *Proceedings of the Section on Statistical Graphics.* Annual Meeting of the American Statistical Association. American Statistical Association, San Francisco.

BOOKSTEIN, F. L. & R. A. REYMENT, 1989. Microevolution in Miocene Brizalina (Foraminifera) studied by canonical variate analysis and analysis of landmarks. *Bulletin of Mathematical Biology,* 51: 657-679.

CHAPMAN, R. E., 1990. Conventional Procrustes approaches. In F. J. Rohlf & F. L. Bookstein (eds.), *Proceedings of the Michigan Morphometrics Workshop.* University of Michigan Museum of Zoology, Ann Arbor.

CRESPI, B. J. & F. L. BOOKSTEIN, 1989. A path-analytic model for the measurement of selection on morphology. *Evolution,* 43: 18-28.

DARROCH, J. N. & J. E. MOSIMANN, 1985. Canonical and principal components of shape. *Biometrika,* 72: 241-252.

EHLINGER, T. J., 1991. Allometry and analysis of morphometric variation in the bluegill, Lepomis macrochirus. *Copeia,* 1991: 347-357.

FELSENSTEIN, J., 1985. Phylogenies and the comparative method. *American Naturalist,* 125: 1-15.

FELSENSTEIN, J., 1988. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics,* 19: 445-471.

FRANCIS, R. C., J. V. BAUMGARTNER, A. C. HAVENS & M. A. BELL, 1986. Historical and ecological sources of variation among lake populations of threespine sticklebacks, *Gasterosteus aculeatus,* near Cook Inlet, Alaska. *Canadian Journal of Zoology,* 64: 2257-2265.

GOODALL, C., 1991. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society,* B 53: 285-339.

HART, P. J. B. & A. B. GILL, 1992. Evolution of foraging behavior in threespine stickleback. In M. A. Bell and S. A. Foster (eds.), *Evolutionary Biology of Threespine Stickleback.* Oxford University Press, Oxford. In press.

HARVEY, P. H. & G. M. MACE, 1982. Comparisons between taxa and adaptive trends: problems of methodology. In K. C. S. Group (eds.), *Current Problems in Sociobiology.* Cambridge University Press, Cambridge.

HARVEY, P. H. & M. D. PAGEL, 1991. *The Comparative Method in Evolutionary Biology.* Oxford University Press, Oxford.

JOLICOEUR, P., 1963. The multivariate generalization of the allometry equation. *Biometrics,* 19: 497-499.

KEAST, A. & D. WEBB, 1966. Mouth and body form relative to feeding ecology in the fish fauna of a small lake, Lake Opinicon, Ontario. *Journal of the Fisheries Research Board of Canada,* 23: 1845-1874.

KUHRY, B. & L. F. MARCUS, 1977. Bivariate linear models in biometry. *Systematic Zoology,* 26: 201-209.

LARSON, G. L., 1976. Social behavior and feeding ability of two phenotypes of Gasterosteus aculeatus in relation to their spatial and trophic segregation in a temperate lake. *Canadian Journal of Zoology,* 54: 107-121.

LAVIN, P. A. & J. D. MCPHAIL, 1985. The evolution of freshwater diversity in the threespine stickleback *(Gasterosteus aculeatus):* site-specific differentiation of trophic morphology. *Canadian Journal of Zoology,* 63: 2632-2638.

LAVIN, P. A. & J. D. MCPHAIL, 1986. Adaptive divergence of trophic phenotype among freshwater populations of the threespine stickleback *(Gasterosteus aculeatus). Canadian Journal of Fisheries and Aquatic Science,* 43: 2455-2463.

LAVIN, P. A. & J. D. MCPHAIL, 1987. Morphological divergence and the organization of trophic characters among lacustrine populations of the threespine stickleback *(Gasterosteus aculeatus). Canadian Journal of Fisheries and Aquatic Science,* 44: 1820-1829.

LELE, S., 1991. Some comments on coordinate-free and scale-invariant methods in morphometrics. *American Journal of Physical Anthropology,* 85: 407-417.

MACLEOD, N. & J. A. KITCHELL, 1990. Morphometric and evolutionary inference: a case study involving ontogenetic and developmental aspects of foraminiferal evolution. In F. J. Rohlf and F. L. Bookstein (eds.), *Proceedings of the Michigan Morphometrics Workshop.* University of Michigan Museum of Zoology, Ann Arbor.

MCARDLE, B. H., 1988. The structural relationship: regression in biology. *Canadian Journal of Zoology,* 66: 2329-2339.

MCPHAIL, J. D., 1984. Ecology and evolution of sympatric sticklebacks (Gasterosteus): morphological and genetic evidence for a species pair in Enos Lake, British Columbia. *Canadian Journal of Zoology,* 62: 1402-1408.

MCPHAIL, J. D., 1992. Speciation and the evolution of reproductive isolation in the sticklebacks *(Gasterosteus)* of southwestern British Columbia. In M. A. Bell and S. Foster (eds.), *Evolutionary Biology of Threespine Stickleback.* Oxford Univeristy Press, Oxford. In Press.

MCPHAIL, J. D. & C. C. LINDSEY, 1970. Freshwater fishes of northwetern Canada and Alaska. *Bulletin of Fisheries Research Board of Canada,* 173: 1-381.

MOSIMANN, J. E. & F. C. JAMES, 1979. New statistical methods for allometry with application to Florida red-winged blackbirds. *Evolution,* 33: 444-459.

OLSHAN, A. F., A. F. SIEGEL & D. R. SWINDLER, 1982. Robust and least-squares orthogonal mapping: methods for the study of cephalofacial form and growth. *American Journal of Physical Anthropology,* 59: 131-137.

PAGEL, M. D. & P. H. HARVEY, 1988. Recent developments in the analysis of comparative data. *Quartely Review of Biology,* 63: 413-440.

RAYNER, J. M. V., 1985. Linear relations in biomechanics: the statistics of scaling functions. *Journal of Zoology,* 206: 415-439.

REILLY, S. M., 1990. Comparative ontogeny of cranial shape in salamanders using resistant fit theta rho analysis. In F. J. Rohlf and F. L. Bookstein (eds.), *Proceedings of the Michigan Morphometrics Workshop.* Special Publication No. 2. The University of Michigan Museum of Zoology, Ann Arbor.

REIMCHEN, T. E., E. M. STINSON & J. S. NELSON, 1985. Multivariate differentiation of parapatric and allopatric populations of threespine stickleback in the Sangan River watershed, Queen Charlotte Islands. *Canadian Journal of Zoology,* 63: 2944-2951.

RICKER, W. E., 1984. Computation and uses of central trend lines. *Canadian Journal of Zoology,* 62: 1897-1905.

RIDLEY, M., 1983. *The Explanation of Organic Diversity: the Comparative Method and Adaptations for Mating.* Oxford University Press, Oxford.

RISKA, B., 1991. Regression models in evolutionary allometry. *American Naturalist,* 138: 283-299.

ROHLF, F. J., 1990. Rotational fit (Procrustes) methods. In F. J. Rohlf and F. L. Bookstein (eds.), *Proceedings of the Michigan Morphometrics Workshop.* University of Michigan Museum of Zoology, Ann Arbor.

ROHLF, F. J. & F. L. BOOKSTEIN, 1990. Proceedings of the Michigan Morphometrics Workshop. University of Michigan Museum of Zoology, Ann Arbor.

ROHLF, F. J. & D. SLICE, 1990. Extensions of the procrustes method for the optimal superimposition of landmarks. *Systematic Zoology,* 39: 40-59.

SEIM, E. & B.-E. SAETHER, 1983. On rethinking allometry: which regression to use? *Journal of Theoretical Biology,* 104: 161-168.

SIEGEL, A. F., 1982. Robust regression using repeated medians. *Biometrika,* 69: 242-244.

SIEGEL, A. F. & R. H. BENSON, 1982. A robust comparison of biological shapes. *Biometrics,* 38: 341-350.

SIEGEL, A. F., 1982. Geometric data anlysis: an interactive graphics program for shape comparisons. In R. L. Launer and A. F. Siegel (eds.), *Modern Data Analysis.* Academic Press, New York.

SOKAL, R. R. & F. J. ROHLF, 1981. *Biometry. The Principals and Practice of Statistics in Biological Research.* W. H. Freeman, San Francisco.

TAYLOR, E. B. & J. D. MCPHAIL, 1986. Prolonged and burst swimming in anadromous and freshwater threespine stickleback, *Gasterosteus aculeatus. Canadian Journal of Zoology,* 64: 416-420.

WEBB, P. W., 1982. Locomotor patterns in the evolution of actinopterygian fishes. *American Zoologist,* 22: 329-342.

WEBB, P. W., 1984. Body form, locomotion and foraging in aquatic vertebrates. *American Zoologist,* 24: 107-120.

WEIHS, D., 1989. Design features and mechanics of axial locomotion in fish. *American Zoologist,* 29: 151-160.

WITHLER, R. E. & J. D. MCPHAIL, 1985. Genetic variability in freshwater and anadromous sticklebacks *(Gasterosteus aculeatus)* of southern British Columbia. *Canadian Journal of Zoology,* 63: 528-533.

LANDMARK DATA: SIZE AND SHAPE ANALYSIS IN SYSTEMATCS. A CASE STUDY ON OLD WORLD TALPIDAE (MAMMALIA, INSECTIVORA)

ANNA LOY & MARCO CORTI

Dept. of Animal and Human Biology
University of Rome 'La Sapienza',
Via Borelli 50 – 00161 Rome, Italy
CORTI@ITCASPUR.BITNET

LESLIE F. MARCUS

Queens College of CUNY, Flushing, New York 11367 & American Museum of Natural History, New York, New York 10024
LAMQC@CUNYVM.BITNET

CONTENTS

ABSTRACT

Morphometric divergence in moles was studied using Bookstein shape coordinates, at both the intraspecific and interspecific level. 13 landmarks were recorded on the right half dorsal view of the skull for 7 fossorial species in the three genera of the family Talpidae, including: Old World *Talpa* (five species); *Mogera* (one species), and New World *Parascalops* (one species). The latter two species were used as outgroups. All coordinates were standardized to a common baseline of unit length. Centroid size was examined for intra–specific and inter–specific variation. Similarities among populations and species in Bookstein shape coordinates were summarized using Mahalanobis D^2 and an UPGMA phenogram. Shape differences among European moles (genus *Talpa*) were further investigated in terms of uniform and non–uniform shape components using landmark based methods.

The UPGMA phenogram is congruent with the systematic hierarchy and phylogenetic hypothesis derived from genetic and cytogenetic data. Size is revealed as the unique component of sexual dimorphism. Size and the uniform component are almost uncorrelated and together represent good descriptors of interspecific variation. A non–uniform component also contributes to our understanding of the phylogenetic relationships among European moles.

INTRODUCTION

Morphometrics on landmark based data has rarely been used in systematics. Up to now, most studies of within and between species variation have been based on 'traditional morphometrics' (sensu Marcus, 1990), e.g. the study of variation and covariation of distance measurements. A few published applications using landmarks include Bookstein and Reyment (1989) on the Miocene foraminifer *Brizalina*, Abe *et al.* (1988) on Cretaceous ostracod *Veenia*, and Tabachnick and Bookstein (1990a, b) on the Miocene foraminifer *Globorotalia*. We believe it is essential to explore the potentiality of the new geometric morphometrics (Rohlf and Bookstein, 1990; Bookstein, 1991) in the description of evolutionary relationships.

These innovative statistical and geometric techniques offer new incites to the study of evolutionary divergence. They provide another tool for determining significant differences and provide new characters for hypotheses of relationship. They also show how anatomy, depicted by the geometric relation between landmarks, differs among populations and taxa of organisms in a way that may be related to past history (phylogeny) and to ecology. These objectives are likely to be achieved through the splitting of morphological variation into its various components, size, and uniform and non–uniform shape (see Bookstein, 1991, and Reyment, 1991, for detailed discussions).

We use shape coordinates to investigate the changes in morphology that occurred during intraspecific and interspecific divergence in some species of insectivorous mammals in the family Talpidae, strictly fossorial mammals. The subterranean habitat is characterized by unique ecological parameters that force microclimatic stability and establish environmental constraints. These in turn support a high degree of specialization and morphological convergence among species (Nevo, 1991). Subterranean animals are thus suitable subjects for a study on how size and various shape components of morphological changes, both linear (uniform or affine, sensu Bookstein, 1990), and non linear (non– uniform or non–affine) landmark changes can be related to phylogeny and adaptation.

Taxa examined include the Palaearctic genus *Talpa*, the Asian genus *Mogera*, and the Neartic genus *Parascalops*. Moles of the family Talpidae are widely distributed in the temperate regions of Holartica (Fig. 1). The systematics of the family is still under revision. We have selected parts of the most recent species

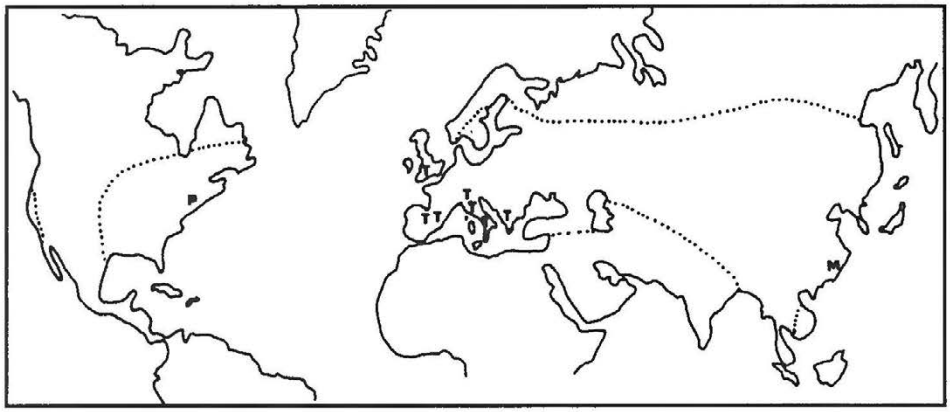


Fig. 1
*World range of the family Talpidae (areas included in dotted lines, from Yates and Moore, 1990), and location of samples: **P** = Parascalops breweri; **M** = Mogera latouchei ; **T** = Talpa sp*

classifications (Honacki *et al.*, 1982; Ramalhinho, 1985; Corti & Loy, 1987; Filippucci *et al.*, 1987; Yates & Moore, 1990; Corbet & Hill, 1991). Samples analyzed represent the following species: Hairy tailed mole (*Parascalops breweri*), occurring in North–East USA and South–East Canada; *Mogera latouchei*, South–East China, Hinan, and Assam, (included in *Talpa* by Corbet and Hill, 1991, but here we use the nomenclature of Yates and Moore, 1990); European mole (*Talpa europaea*), Europe, West Siberia; Roman mole (*T. romana*), Central and Southern Italy, *T. stankovici*, Balkans; Mediterranean mole (*T. caeca*), South Europe and Caucasus; *T. occidentalis*, Spain and Portugal.

Yates and Moore (1990) propose a phylogeny of the Talpidae on the basis of genic, cytogenetic and morphologic characters. They suggest that the *Talpa* and *Mogera* form a monophyletic group. *Parascalops* is not a sister to this group, but has been placed in a clade in a different part of the phylogeny, which appears at present to be paraphyletic.

MATERIAL & METHODS

We examined a total of 113 individuals in 10 samples. They represent three populations of *T. europaea*, from Northern, Western, and Central–South parts of the species range (from England, Spain and Switzerland respectively), two populations of *T. romana*, from the type locality (Lazio, Italy) and from a Southern locality (Calabria, Italy); and one population each for the other 5 species (Fig. 1, Table 1). Specimens are housed in various institutions in Europe and the United States of America (Table 1).

Table 1

Sample localities and collections: MAC = Museo di Anatomia Comparata, University of Rome 'La Sapienza'; VER = Museo di Storia Naturale di Verona (Verona, Italy); MAD = Museo Nacional de Ciencias Naturales de Madrid (Madrid, Spain); BRI = British Museum of Natural History; AMNH = American Museum of Natural History (New York, USA); BER = Museum d'Histoire Naturelle Berne (Berne, Switzerland)

SPECIES	LOCALITY	COLL.	CODE	M	F	TOT
Talpa europaea	Avendano (Pyrenee, Spain)	MAD	TEUSP	6	11	17
Talpa europaea	Berne (Switzerland)	BER	TEUBE	13	14	27
Talpa europaea	Galles (Great Britain)	BRI	TEUEN	7	4	11
Talpa romana	Ostia (Lazio, Italy)	MAC	TRO	14	10	24
Talpa romana	Fiumefreddo (Calabria, Italy)	MAC	TROCA	4	1	5
Talpa stankovici	M. Vitzi (Macedonia, Greece)	MAC	TST	1	2	3
Talpa caeca	Zumaglia (Italy)	VER	TCA	3	3	6
Talpa occidentalis	Salamanca (Spain)	MAD	TOC	3	3	6
Mogera latouchei	Chung Hsian (Fuxien, China)	AMNH	MIN	4	1	5
Parascalops breweri	Wayre (Pennsylvania, USA)	AMNH	PBR	3	6	9
Total				58	55	113

We photographed the dorsal view of each skull using a Nikon FE camera equipped with a 50mm macro lens and a circular flash. The focal plane of the camera was parallel to the dorsal surface of the skull, and centered on bregma. Pictures were enlarged x 4 on glossy paper. Landmarks were collected on a Calcomp 2200 digitizing tablet using a modification of Lessoft caliper software (Marcus, 1988).

We collected 13 landmarks on the right half of the skull, to avoid the effect of lateral asymmetry (Fig. 2). The landmarks are points at the: 1 – tip of the rostrum, 2 – maximum width of the rostrum (point of maximum curvature), 3 – minimum width of the rostrum (point of flexus); 4 – maximum width of palatine (point of maximum curvature); 5 – anterior inner curvature of zygomatic arch (point of flexus); 6 – distal extremity of coronal suture; 7 – posterior inner curvature of zygomatic arch (point of flexus); 8 – maximum width of the bulla (point of maximum curvature); 9 – asterion; 10 – tip of postparietal; 11 – lambda; 12 – bregma; 13 – nasion. These are landmarks of different types according to the classification of Bookstein (1991): landmarks 9, 11, and 12 are type one; 6, 8, and 13 are type two; and 1, 2, 3, 4, 5, 7, and 10 are type three.

Error of measurements was evaluated by computing the standard deviation of measurements from 3 separate digitizing sessions at 20 days intervals) on a subsample of 39 individuals. Before all further computation raw coordinates were rotated and translated to have the X axis described by landmarks 11 and 13, with

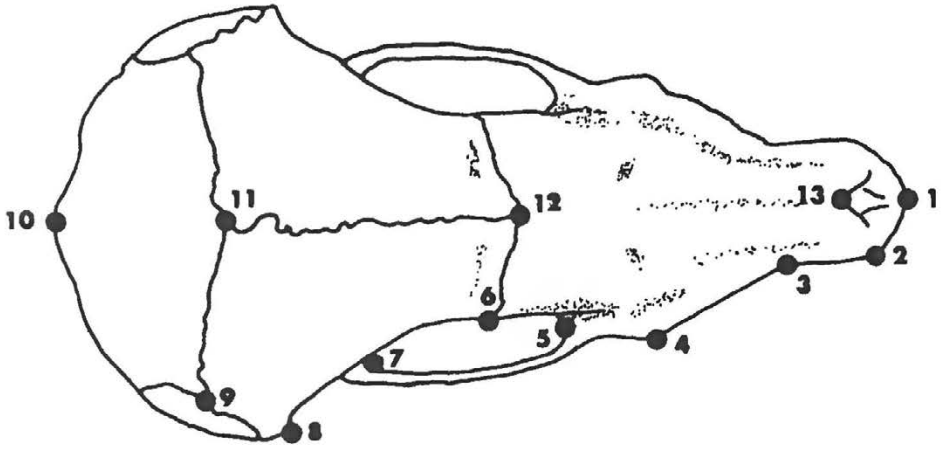


Fig. 2

Location of landmarks for which coordinate data was collected on the dorsal view of each skull

the origin located on landmark 11, using another modification of Lessoft caliper software (Marcus, 1988).

Baseline endpoints (Bookstein, 1990, 1991) were selected from the five landmarks along the midline of the skull (numbers 1, 10, 11, 12, 13 in Fig. 2). Two alternative baselines were evaluated. One baseline was between landmarks 1 and 10 (proximal and distal tips of the skull), and the other between 11 and 13 (lambda and nasion respectively). Baseline 11 – 13 was almost as long and was preferred as it appeared to introduce less distortion in the alignment of the skulls along the midline. The 'midline' points did not always lie strictly on a straight line, especially landmark 10. The length of the baseline and centroid size were compared.

All of the coordinates were rotated to the baseline and divided by its length to produce Bookstein shape coordinates (Mardia and Dryden, 1989) using the SAS IML program UNICOORD written by L.F. Marcus (corrected and modified from version in appendix in Reyment, 1991). Centroid size, as the square root of the mean squared distance from each point to the centroid of an object, was determined as well. Bookstein shape coordinates have a baseline with one end at (0,0) and the other at (1,0), therefore the baseline length is 1. Bookstein shape coordinates thus describe the position of the non–baseline landmarks as vertices of triangles having the baseline as a common side. Similarities and differences in shape may then be studied in terms of the distribution or scatter of these vertices.

Baseline size and centroid size were compared by bivariate scatter and correlation. Analysis of variance of centroid size included: two way analysis (unbalanced design) to test for sexual dimorphism, population differences, and sex–population interaction using PROC GLM in SAS. Since there was no interaction or sexual

dimorphism detected, a one way analysis of variance was subsequently used for group comparisons.

Univariate and multivariate analyses of the shape coordinates provided estimates of differences among populations. Canonical variate analysis and Multivariate Analysis of Variance (MANOVA) were used to examine and test for population differences for the multivariate assemblage of coordinates.

Differences among populations were summarized by Mahalanobis D^2 and these were clustered in a phenogram using the Unweighted Pair Group Method (UPGMA) in the program package NTSYS–pc (Rohlf, 1990). Significance of differences between group means was tested using Hotelling's T^2 statistic and significance was adjusted for multiple comparisons by use of the Bonferroni inequality (Marcus, 1990). A minimum spanning tree was found for Mahalanobis D's and superimposed on a bivariate plot of the means of the first two canonical variates using NTSYS.

Uniform X and Y factors were extracted from the Bookstein shape coordinates following the formula in Bookstein (1990), and as part of the output of UNICOORD. These factors were regressed on size and on each other.

The coordinates were also analyzed using superimposition methods, which partition shape differences into linear (uniform or affine) and non linear (non–uniform or non–affine) components. For each sample the Generalized Least Square option (GLS in **GRF** software, Rohlf and Slice, 1990), was used to find consensus configurations for the specimens. Bivariate plots of landmark means for each population were produced and overlain on one plot. Deviation vectors were plotted for each population from other populations for intraspecific comparisons, from *Talpa europaea* for species of *Talpa*, and from *Parascalops breweri* consensus for all species of *Talpa*.

The consensus configurations for each of the five species of genus *Talpa* were then compared using the method of relative Warps and the Thin Plate Spline Relative Warp program (**TPSRW**, Rohlf, 1990; version May 1992).

Relative Warps were determined with respect to the mean configuration for the five species as a reference and using the option alpha= 0. Rohlf, this volume, suggests this option in exploratory systematic studies as warps are equally weighted. A choice of α=1 emphasizes large scale non–linear deformations over local or small–scale ones, and this strategy appears to be more appropriate in ontogenetic or evolutionary sequence comparisons. The minimum energy superposition option was used, and weights were based on raw data as deviations from means. Relative warp scores were computed both retaining and removing the affine eigenvectors, and the matrices produced were compared using the Mantel test in NTSYS–pc.

Finally, to test for relations among components of morphometric variation, results from different analyses were compared. The uniform factor, relative warps 1, 2 and 3, were regressed on centroid size.

RESULTS

Measurement error

Standard deviation varied between 0.002 and 0.019 over the 13 landmarks, and the average standard deviation was 0.009. This corresponds to 13 digitising units, which in turn translates to 0.3 mm on a skull of about 30 mm. We considered this error to be negligible.

Sexual dimorphism

Analysis of Variance did not show significant interaction, that is differences in dimorphism among populations, or a sexual dimorphism main effect in shape coordinates (Table 2). However multivariate analysis of variance found a significant interaction for shape coordinates for only Roy's largest root criterion. An analysis of variance of the canonical variate corresponding to this root, also showed highly significant interaction. The coefficients for this variate were a complicated function of all of the coordinates, and showed no simple interpretable relation among the coordinates.

Sexual dimorphism is strongly present for centroid size (Table 4) and the logarithm of centroid size. There is no interaction between population and sex, except for the exception pointed out above, so that the results are interpreted as isometric differences between males and females: males are bigger than females

Table 2

*Two way analysis of variance testing sexual dimorphism against population variation for all shape coordinates; and one way analysis of variance combining sexes. Mean squares from SAS Type III GLM procedure for unbalanced designs. All mean squares should be multiplied by 10^{-4}. *alpha=.05; ** 0.01; *** 0.001.*

Analysis of Variance Tables

	Two Way Analysis				One Way Analysis	
	POP.	SEX	POP. * SEX	ERROR	POP.	ERROR
D.F.	9	1	9	93	9	103
COORD.						
X1	40.86***	0.50	0.87	1.48	45.32***	1.41
Y1	1.01**	0.00	0.46	0.35	0.91*	0.35
X2	16.50***	0.09	0.79	1.60	17.35***	1.51
Y2	8.93***	0.01	.73	0.76	9.51***	0.75
X3	17.70***	0.04	1.06	2.62	18.85***	2.47

Table 2 (Continued)

	Two Way Analysis				One Way Analysis	
	POP.	SEX	POP. * SEX	ERROR	POP.	ERROR
Y3	35.34***	0.01	0.59	0.70	5.34***	0.69
X4	51.65***	0.91	1.16	2.91	53.59***	2.73
Y4	13.79***	0.29	1.28	1.22	14.04***	1.22
X5	52.70***	1.09	0.81	2.58	55.32***	2.44
Y5	13.88***	0.74	0.76	1.04	14.59***	1.01
X6	54.19***	0.32	10.00	8.52	66.56***	8.58
Y6	11.61***	0.38	1.71	1.33	12.86***	1.36
X7	62.87***	0.15	1.30	4.42	65.51***	4.12
Y7	13.08***	2.28	1.18	1.58	14.39***	1.54
X8	122.57***	1.23	1.81	8.68	129.95***	8.00
Y8	46.30***	0.21	2.32	3.17	48.40***	3.07
X9	59.34***	6.66	5.12	9.24	63.68***	8.82
Y9	85.34***	0.22	2.77	3.80	90.86***	3.68
X10	317.15***	17.22	5.75	12.21	333.13***	11.66
Y10	11.56**	0.03	4.79	3.81	10.56**	3.89
X12	108.36***	0.02	10.21	1.47	116.20***	13.24
Y12	1.30	0.83	0.67	0.84	1.78*	0.82

Table 3

*Results from multivariate analysis of variance testing sexual dimorphism against population variation for all shape coordinates, p values, SAS GLM Type III Sums of Squares. Note interaction for Roy's Greatest Root significant vector is:(–18*x1–96*y1–34*x2+107*y2+11*x3 +12*y3+16*x4–17*y4–30*x5–27*y5–31*x6+9*y6+40*x7+50*y7–27*x8–30*y8 +4*x9–40*y9+21*x10+28*y10+6*x12–61*y12).* alpha=0.05; ** .01; *** .001*

Source	Wilk's Lambda Num,Den DF	Pillai's Trace Num,Den DF	Hotelling–Lawley Trace Num,Den DF	Roy's Greatest Root Num,Den DF
pop.	8.45***	5.30***	13.04*** 198,632	65.81***
	198,621.5	198,720		22,80
sex	0.61	0.61	0.61	0.61
	22,72	22,72	22,72	22,72
pop.* sex	0.75	0.76	0.75	2.20**
	198,621.5	198,720	198,632	198,632

by the same amount for each population (Table 5) or by the same proportion for the logged data. The small sample sizes account for the similarities.

These results justified combining the shape coordinate data over sexes for each species and ignoring sex in further analyses. This made the analysis simpler because of the disparate sample sizes for sexes among populations.

Table 4
Results from two way analysis of variance of centroid size. alpha=.001 ***

	DF	Mean Squares	F
population	9	180358	95.91***
sex	1	23979	12.75***
pop. * sex	9	1249	0.74
error	93	1880	

Population comparisons: distances

As it is shown in Table 5, size is also an important component of population and species differences, and must be considered both by itself and together with shape differences. Figure 3 shows the relation between centroid size and baseline size for all populations. This plot demonstrates the clear differences in size among

Table 5
Centroid size in digitizing units for males and females of each population

Centroid Size POP	SEX	Mean	Overall Mean
TEUBE	M	1601	1572
	F	1545	
TEUSP	M	1683	1637
	F	1612	
TEUEN	M	1588	1573
	F	1548	
TRO	M	1823	1797
	F	1760	
TROCA	M	1677	1663
	F	1610	
TST	M	1670	1635
	F	1618	
TCA	M	1439	1440
	F	1440	
TOC	M	1491	1480
	F	1468	
MIN	M	1275	1274
	F	1267	
PBR	M	1496	1480
	F	1472	

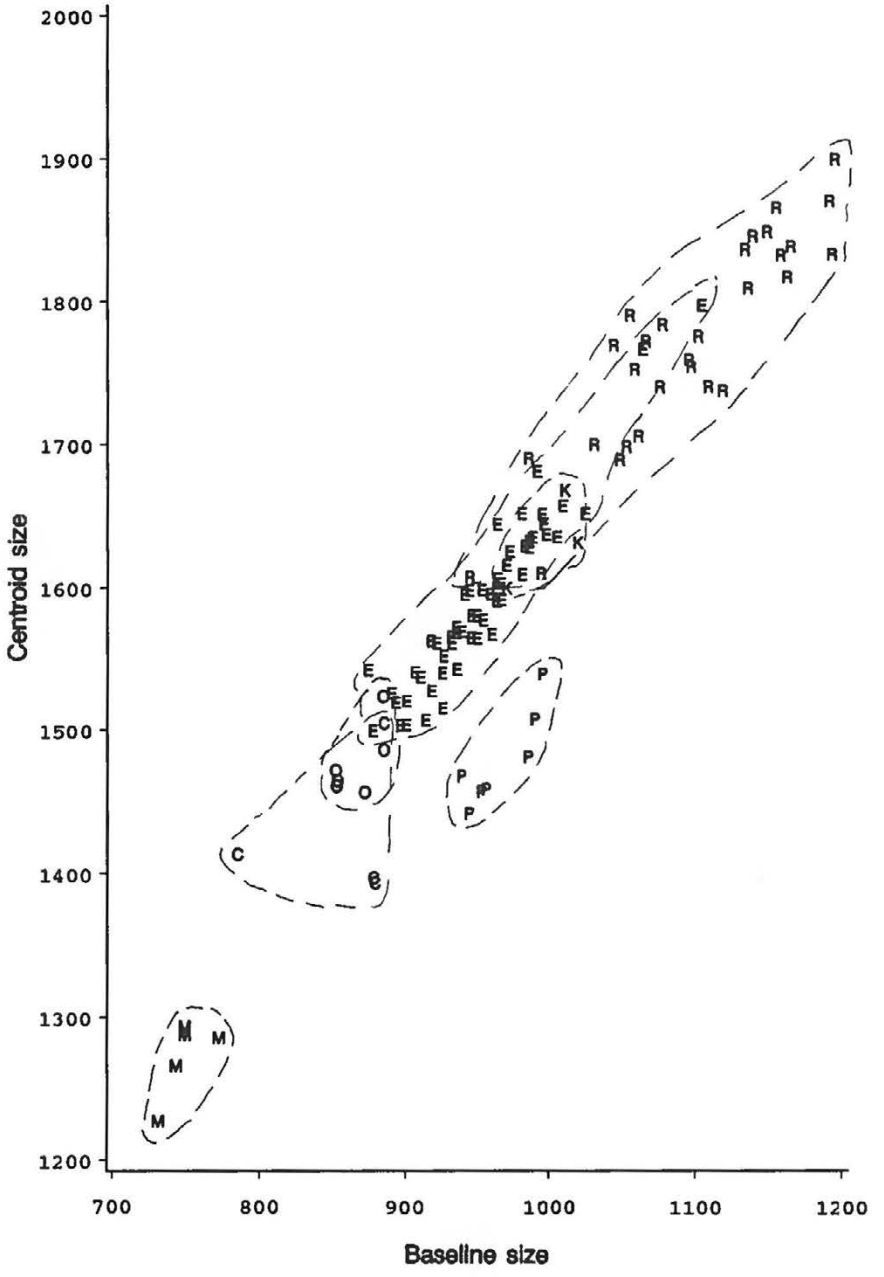


Fig. 3

Scatter–plot of individuals showing relation between baseline and centroid size for all specimens. Baseline "size" is the distance from landmark 11 to 13 in digitizing units. Range of variation for each species enclosed by dashed lines

P = *Parascalops breweri*; **M** = *Mogera latouchei;* **E** = *Talpa europaea*; **R** = Talpa *romana*; **K** = *Talpa stankovici*; **C** = *Talpa caeca*; **O** = *Talpa occidentalis*

species, and particularly the separation between the 'small moles' *Talpa caeca* and *Talpa occidentalis* and the 'big moles' *Talpa romana* and *Talpa stankovici*, a fact already known and well described in the literature (Toschi and Lanza, 1959; Capolongo and Panascì, 1976; Capanna, 1981; Corti and Loy, 1987; Gonzales and Roman, 1989). It is also interesting to notice that in *Parascalops* there is a different covariance between centroid size and baseline size.

19 of the 22 Bookstein shape coordinates are highly significantly different among populations (Table 2, last column). The high discriminatory power of these characters is confirmed by multivariate analysis: Mahalanobis D^2 is always large, even for intraspecific differences (Table 6), and the F – value is always significant (0.05 level) for each paired comparison, and most are highly significant. Bonferroni adjustment for the comparison of 10 samples, requires a 0.05/45 probability level for each comparison for an overall 0.05 significance level. This stringent criterion is satisfied by all of the comparisons. No overlap was observed between groups, with 100% correct assignment for all samples, though this is not unusual for such small samples with so many variables.

Table 6

Mahalanobis distances squared computed on canonical variate centroids of all Bookstein coordinates for all populations. D^2 values above diagonal and unbiased D^2 below the diagonal. All values significant at 0.0001 level. See table one for population codes

	PBR	MIN	TCAE	TEUBN	TEUEP	TEUS	TOCA	TROC	TRO	TST
PBR	0	103	132	142	112	155	85	109	70	111
MIN	90	0	127	87	39	85	46	89	92	130
TCA	117	111	0	83	90	102	45	55	91	100
TEUBE	127	76	73	0	32	34	67	58	86	120
TEUEN	99	33	79	27	0	31	47	72	84	136
TEUSP	139	74	90	30	26	0	79	77	107	184
TOC	74	38	37	59	40	61	0	34	41	59
TROCA	95	77	46	50	62	67	27	0	28	66
TRO	62	81	81	77	75	96	35	23	0	42
TST	97	113	86	105	119	163	49	55	35	0

Mahalanobis distances between centroids are congruent with the current systematic hierarchy: intraspecific distances are smaller then interspecific ones, and species of the same genus or belonging to the same clade are more similar to each other then outgroup species.

This pattern is evident also from the minimum spanning tree superimposed on the scatter plot for the first two canonical variate means (Fig. 4). *Parascalops* is farthest from all other populations, which is congruent with its distant relationship (Yates and Moore, 1990). *Mogera latouchei* has as its nearest neighbor *Talpa europaea*. The latter is then connected to the other four species of the genus *Talpa*, e.g. the European species that are endemic for restricted areas (*T. romana, T. stankovici, T. occidentalis*), or have a limited distribution (*T. caeca*). Similar relationships are also shown by the UPGMA phenogram computed from Mahalanobis D^2 (Fig. 5): *Parascalops breweri* is most separated from all other species, and therefore may be considered as a morphometric outgroup. *Mogera latouchei* clusters with *Talpa europaea*, while another cluster includes the other species of *Talpa* and shows more similarities among the three endemic species *Talpa romana, T. stankovici* and *T. occidentalis*, with *Talpa caeca* joining this group at a greater morphometric distance.

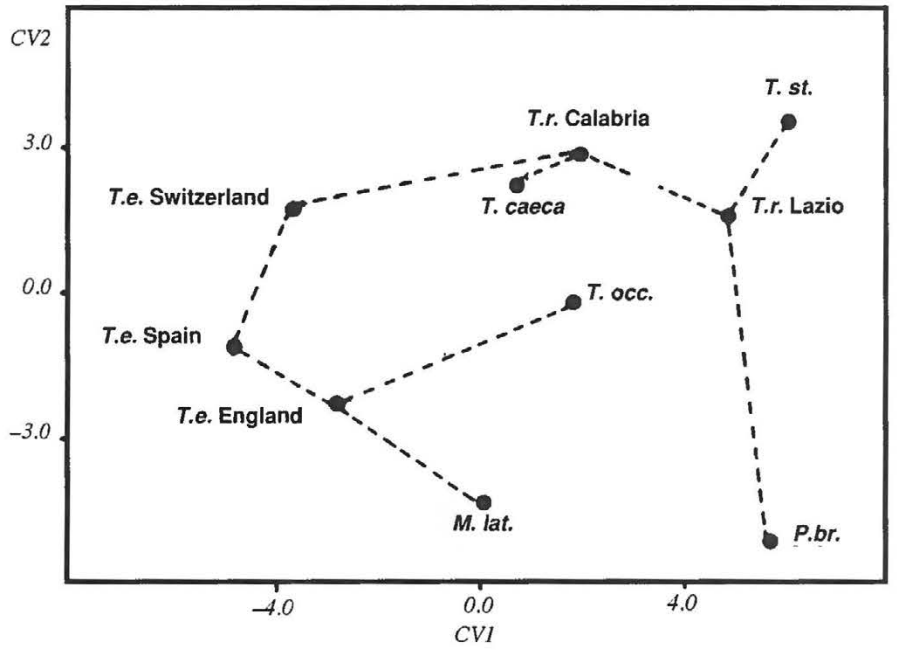


Fig. 4

Minimum spanning tree of sample centroids superimposed on the plot of the first two canonical variates. All 22 Bookstein shape coordinates were used in the analysis of the 10 samples: 3 of *Talpa europaea*, 2 of *Talpa romani* and one each of *Talpa caeca, Talpa occidentalis, Talpa stankovici, Mogera latouchei*, and *Parascalops breweri*

Landmark shape comparisons

To show how landmarks vary in relation to the hierarchical pattern depicted by the UPGMA (Fig. 5), intraspecific and interspecific shape differences were

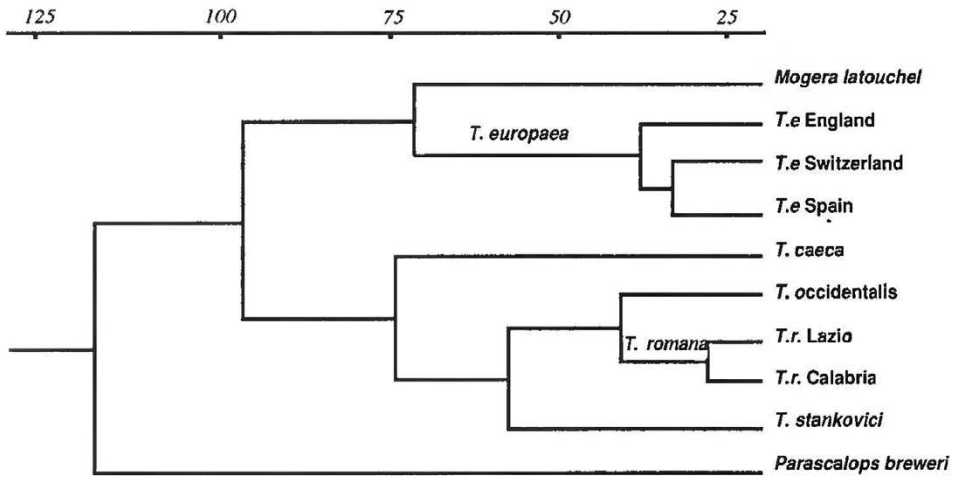


Fig. 5
UPGMA phenogram derived from Mahalanobis squared distances computed on all 22 Bookstein coordinates for all 10 samples

first visualized by superimposing landmarks of consensus specimens (Fig. 6) for each sample. Figure 6A shows the two populations of *T. romana*, expressing the Calabria sample as a deviation from the type Lazio sample, and Figure 6B shows landmark deviations of each population consensus from the other for the three populations of *Talpa europaea*. Directions of landmark deviations differ in the two species, e.g. intraspecific variation involves different landmarks.

Landmark deviations of consensus specimens for the four limited–range species relative to *Talpa europaea* are also heterogeneous (Fig. 6C): most landmarks "move" in opposite directions (e.g. landmarks 3, 6, 8, 9, 10), and *T. europaea* behaves as an 'intermediate' form. Figure 6D shows deviations of consensus configurations of all species of *Talpa* from the reference configuration of *Parascalops breweri*. The difference in shape is clear and the *Talpa* species deviate from *Parascalops* in similar ways generally. There is a general tendency for a wider and shorter braincase and for a longer and wider rostrum.

Figure 7 and Table 7 report results of relative warp analysis run on the five species of *Talpa* superimposed on the average specimen. As can be seen from the table and figure, major modifications are described by horizontal shifting of landmarks, with the X component of principal warp 2, 8, and 10 characterizing the larger part of the variation. These principal warps are influenced by different regions of the skull: principal warp 2 has highest coefficients for landmarks 1, 2, 3 and 13 which describe the rostrum (Fig. 2), and for landmark 6 located on the zygomatic region; principal warp 8, 9 and 10 are more related to the

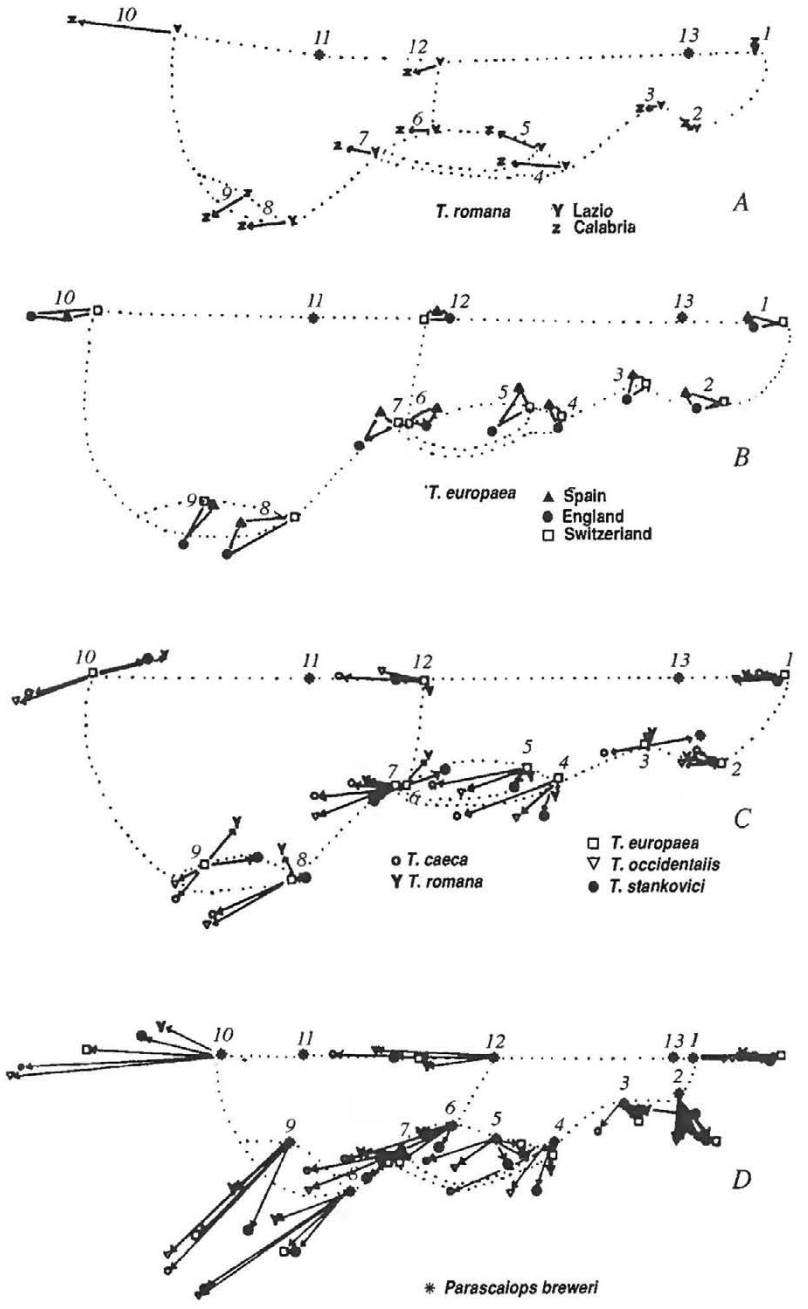


Fig. 6

Superimposition of average consensus configurations and directions of differences for each landmark. Intraspecific variation (deviations of populations from each other): A. T. romana; B. T. europaea; C. landmark deviations of the four limited – range species Talpa from T. europaea; D. deviations of all euroasiatic Talpa samples from Parascalops breweri

braincase (landmarks 9, 10, 11), zygomatic region (landmarks 5 and 6), and the palatine (landmark 4). This general trend of shape changes indicates that species differ for relative lengthening of regions of the skull rather then for their relative width.

Table 7

Results of Relative Warp analysis run on five species of European moles considering the mean configuration as the reference configuration (options: alpha=0; minimum energy as superimposition criterium; deviations from reference; affine eigenvectors not retained). Are reported Principal Warps that have most influence on each Relative Warp, together with main direction of vector changes for each Principal Warp (X or Y) and landmarks that are most related with each Principal Warp

Rel.Warp	Eigenvalue	% Variance	Cumulative	Prin.Warp	Landmark
1	0.02369	39.3	39.3	8X	3–6–10–11
				9X	4–10
2	0.01654	27.4	66.6	2X	1–2–3–6–13
				10X	5–6–9–10–12
3	0.01365	22.5	89.1	8X	3–6–10–11
				10X	5–6–9–10–12

The three dimensional plot of relative warp scores (Fig. 8) shows an interesting species distribution pattern: on the first relative warp axis (39.3% of total square root variance, affine removed) the two extremes are represented by *T. romana* (negative values) and *T. caeca* (positive values), which differ in relative lengthening of the rostral and zygomatic regions; on the second relative warp (27.4% of variance explained) the extremes are represented by *T. europaea* (negative values) and *T. stankovici* (positive values) , which differ mainly by the extension (Fig. 7); and on the third axis (22.5% of variance explained) *T. caeca* is found at the negative extreme and *T. occidentalis* at the positive one, diverging essentially by relative extension of the braincase and the zygomatic region. In the UPGMA phenogram computed from relative warp scores (Fig. 9) *T. romana* and *T. stankovici*, and *T. caeca* and *T. occidentalis* form separate clusters, and these two clusters then join *T. europaea*.

Retention of the affine eigenvectors in the computation of relative warp scores do not affect this pattern of relations (Mantel test for the two matrices gives a value of $r = 0.96$ and $P < 0.003$), but distances between species increase with the retention of the affine eigenvectors (Fig. 10). This is to be expected for $\alpha=0$, and is due to the fact that only three relative warps were retained.

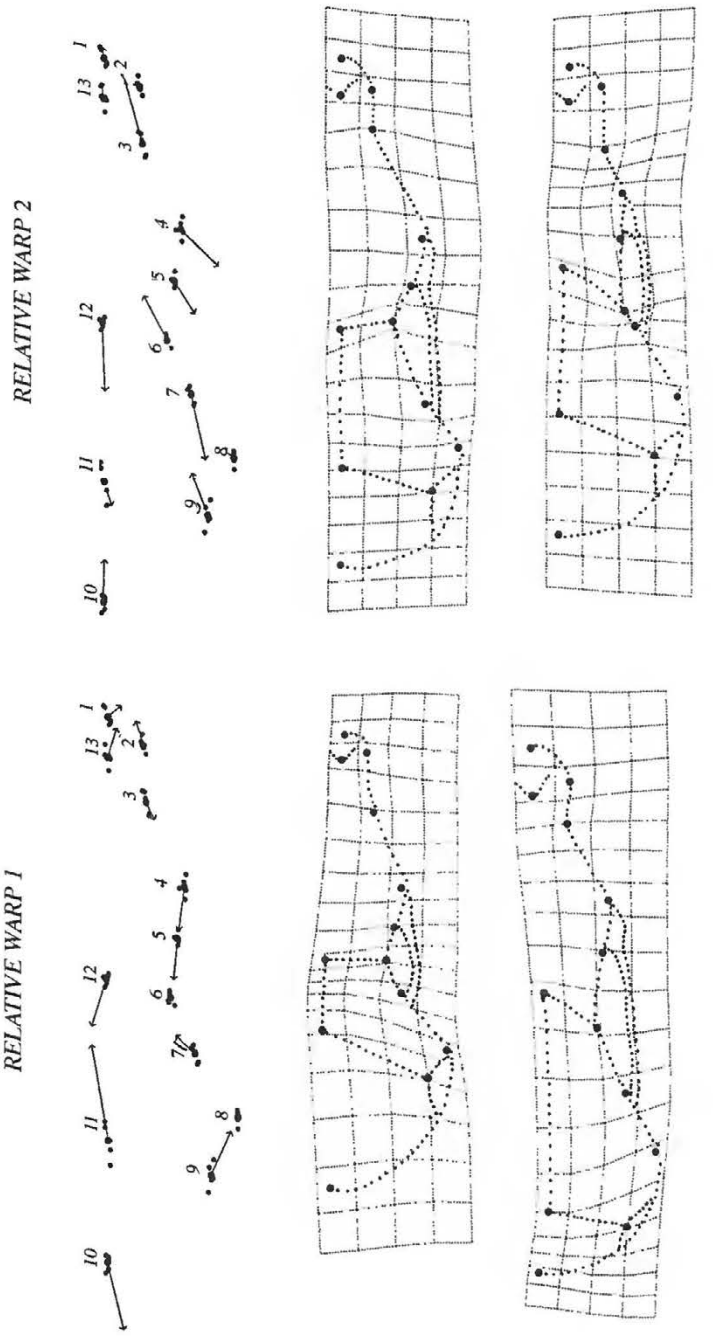


Fig. 7

Vectors of landmark changes and deformation grids for the first and second relative warp computed for comparison among 5 species of the genus Talpa, considering the mean configuration as the reference configuration. Relative warp analysis with $\alpha=0$

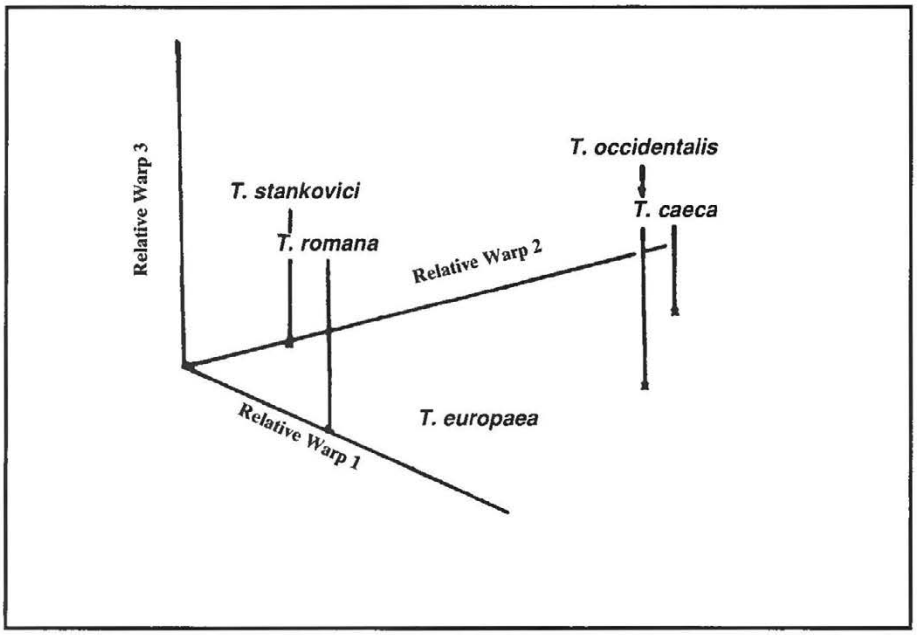


Fig. 8
Top, scatter plot of first three relative warps for the five species of European moles (genus Talpa)

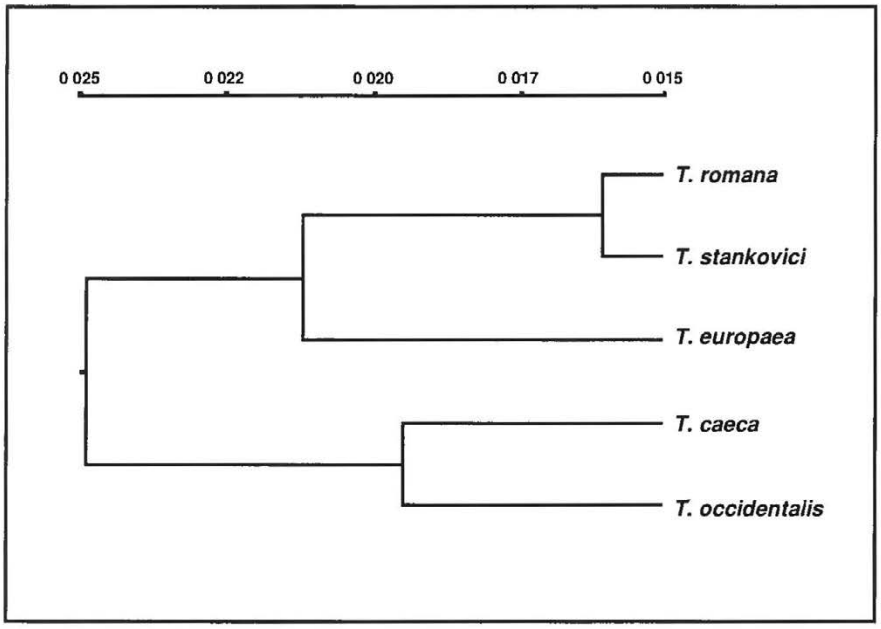


Fig. 9
UPGMA phenogram computed on Euclidean distances in the relative warp space

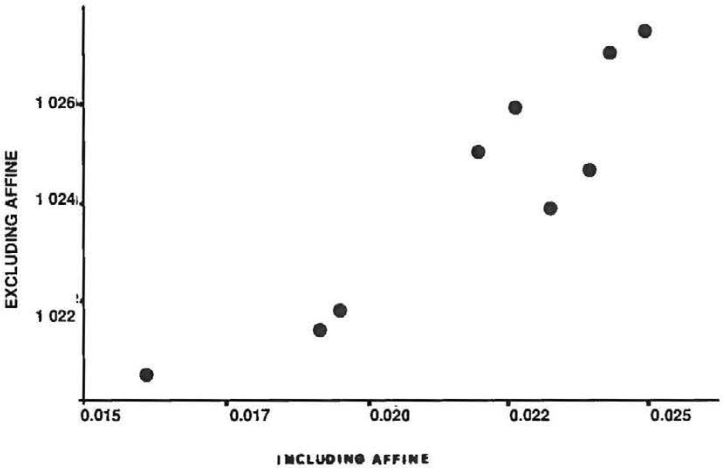


Fig. 10
Mantel plot for Euclidean distances derived from relative warp scores computed both including and excluding the affine component

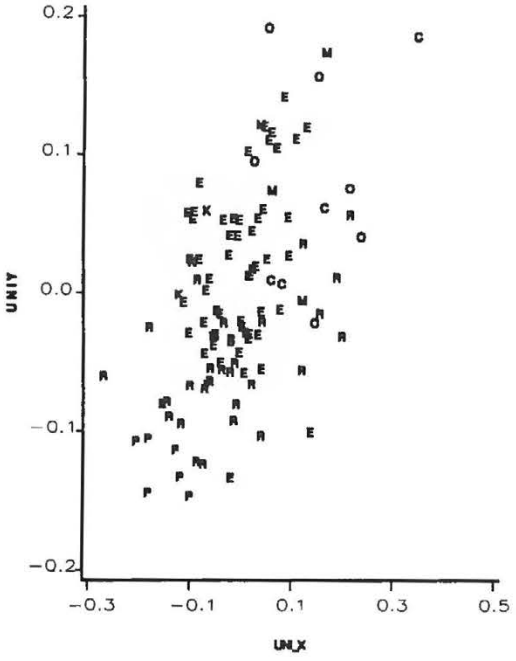


Fig. 11
Scatter plot of Uniform X and Uniform Y for all specimens computed from all 22 Bookstein coordinates **P** = *Parascalops breweri*; **M** = *Mogera latouchei;* **E** = *Talpa europaea*; **R** = *Talpa romana*; **K** = *Talpa stankovici*; **C** = *Talpa caeca*; **O** = *Talpa occidentalis.*

The analysis of the affine (uniform) factor for all samples shows that its components along the X and Y axes are only moderately correlated (Fig. 11) with an r^2 of 0.0336. Figure 11 shows that *Parascalops breweri* is well separated for affine components, with much more overlap among the other species.

Moreover, the uniform components are only slightly correlated with centroid size (Fig. 12, Table 8). The affine factors of size and uniform shape more clearly define the species than centroid size (Fig. 3) or uniform shape differences (Fig. 11) alone.

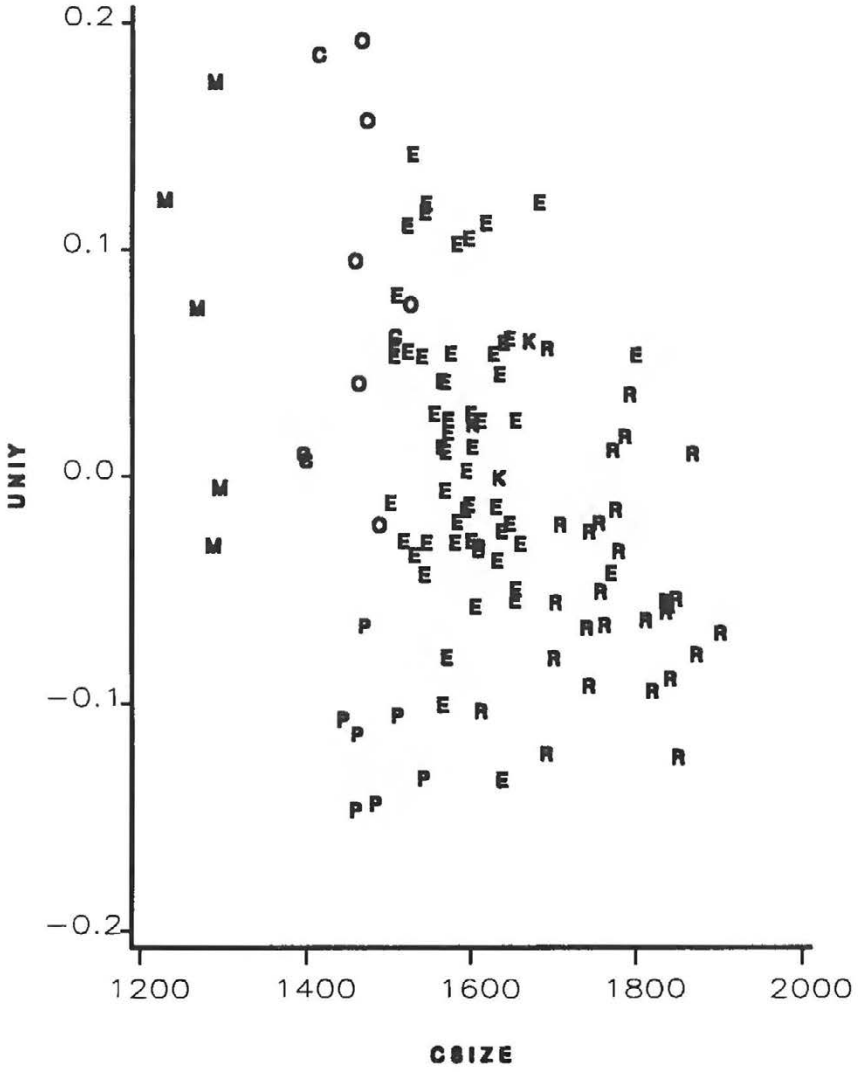


Fig. 12

Scatter plot of Uniform Y factor and Centroid Size for all specimens. **P** = *Parascalops breweri*; **M** = *Mogera latouchei;* **E** = *Talpa europaea*; **R** = *Talpa romana*; **K** = *Talpa stankovici*; **C** = *Talpa caeca*; **O** = *Talpa occidentalis*

The first three relative warps are also not correlated with centroid size (Table 8).

These results offer clear evidence that differences in size among the species are not related to linear or non-linear components of shape variation among the landmarks.

Table 8

Correlation between size (expressed by centroid size) and other components of shape variation, e.g. affine (expressed by Uniform X and Y) and non – affine (expressed by Relative Warps). Uniform factors are estimated for all specimens, while Relative Warps were computed for the five consensus specimens of genus Talpa

	Centroid size
Uniform X	$r^2 = 0.071$
Uniform Y	$r^2 = 0.121$
Relative Warp 1	$r^2 = 0.005$
Relative Warp 2	$r^2 = 0.095$
Relative Warp 3	$r^2 = 0.190$

DISCUSSION

Phenetic distances computed on Bookstein coordinates confirm that *Parascalops breweri* is morphologically distinct from the Old World Talpidae, while *Mogera* is not separated from the *Talpa* group. *Mogera* appears to be more closely related to *T. europaea* then the latter is to the other species of *Talpa* analyzed. The inclusion of *Mogera* in the genus *Talpa* proposed by Corbet & Hill (1991) appears more acceptable in the light of our results. The analysis also demonstrated affinities between south European moles *T. occidentalis, T. stankovici*, *T. romana* and *T. caeca*, which appear to be more closely related to each other then to *T. europaea*. This suggests that the latter belongs to a different ancestral stock.

Separate analysis of different components of variation, e.g. size, affine, and non–affine variation helped us to clarify relationships among populations and species. Size has been revealed to be the only identifiable component of sexual dimorphism in our data for the dorsal view of the skull, considered in all populations and species of Talpidae included in our study. Moreover, the relation between centroid size and baseline size is the same for all species belonging to the monophyletic group *Talpa* together with *Mogera*, while a different relationship was observed between the two size measures in *Parascalops*. The relation between baseline length and centroid size might represent a synapomorphy, therefore supporting a phylogenetic relationship within the family Talpidae. Of course this

hypothesis should be tested by including in the analyses other sister taxa to *Parascalops*, namely *Scalopus, Scapanus*, and *Scapanulus.*

Non–affine changes in shape of the skull among European species mainly involve horizontal shifting of landmarks that express themselves in different anterior–posterior lengthenings of the braincase, the zygomatic region, the palatine, and the rostrum, rather then in the relative width of these regions.

The UPGMA phenogram derived from the non–affine components of shape changes shows affinities between *T. stankovici* and *T. romana,* and between *T. occidentalis* and *T. caeca*, while *T. europaea* is still separate from these four species. The affine or linear component of shape variation among this group of species is not correlated with size.

Therefore the differences in size among species is not closely linked to linear shape modifications.

Similarly, non affine variation is independent of size as determined from the relations to the relative warp scores. It is important to emphasize the choice of the value 0 for the alpha parameter in the relative warp analysis. A choice of $\alpha=1$ weights the warps so that large–scale deformations are considered more important than small ones. In a taxonomic study such as this, there is no *a priori* reason to do that, and $\alpha=0$ provides an equal weighting option. In a comparison to a putative ancestor in an evolutionary study, or analysis of an ontogenetic series, setting $\alpha=1$ might be preferable.

The non–affine component of shape variation allowed us to detect differences between different parts of the skull that were 'hidden' as convergence when the 'traditional' approach was used. If similarities shown by the non–affine component reflects common historical events, one may hypothesize a common origin of *T. romana* and *T. stankovici*, dating back to the earliest Pleistocene glacial events (Donau, Gunz), when the land link between the Balkans and Italy was established.

ACKNOWLEDGMENTS

We wish to extend our gratitude to Elisa Bello and to Antonio Valdecasas for their friendly assistance during the Valsain Workshop. We thank Prof. E. Capanna for the generous support he has given to our work. We would like to thank the curators of American Museum of Natural History, British Museum (Natural History), Museum d'Histoire Naturelle de Berne, Museo di Storia Naturale di Verona, Museo di Anatomia Comparata dell'Universita' di Roma, and Museum de Ciencias Naturales de Madrid for their assistance during our visits. Fred Bookstein offered helpful comments that improved our work. The work was supported by M.U.R.S.T. 40% and by C.N.R. project N. 91.00655.CT04.

REFERENCES

ABE K, R. REYMENT, F. L. BOOKSTEIN, A. HONIGSTEIN, O. HERMALIN, 1988. Microevolutionary changes in *Veenia fawarensis* (Ostracoda, Crustacea) from the Cretaceous (Santonian) of Israel. *Hist. Biol.,* 1: 303–322.

BOOKSTEIN F.L., 1982. Foundations of morphometrics. *Annual Review of Ecology and Systematics,* 13: 451–470.

BOOKSTEIN F. L., 1984. Tensor biometrics for changes in cranial shape. *Annals of Human Biology,* 11: 413–437.

BOOKSTEIN F. L., 1986. Size and shape spaces for landmark data in two dimensions. *Statistical Science,* 1: 181–242.

BOOKSTEIN F. L. 1987. Describing a craniofacial anomaly: Finite elements and the biometrics of landmark location. *American Journal of Physical Anthropology,* 74: 495–509.

BOOKSTEIN F. L., 1989. Principal warps: Thin–plate splines and the decomposition of deformations. *I.E.E.E. Transactions on Pattern Analysis and Machine Intelligence,* 11: 567–585.

BOOKSTEIN F. L., 1990a. Introduction to methods for landmark data. In: *Proceedings of the Michigan Morphometrics Workshop.* Rohlf F.J. & F.L. Bookstein eds. The University of Michigan Museum of Zoology. Special Publication. No.2: Ann Arbor.

BOOKSTEIN F. L., 1990b. Higher order features of shape change for landmark data. In: *Proceedings of the Michigan Morphometrics Workshop.* Rohlf F.J. & F.L. Bookstein eds. The University of Michigan Museum of Zoology. Special Publication No.2, Ann Arbor.

BOOKSTEIN, F. L., 1991. *Morphometric tools for landmark data: geometry and biology.* Cambridge Univ. Press, Cambridge.

BOOKSTEIN, F. L., B. CHERNOFF, J.M. HUMPHRIES JR., G.R. SMITH & R.E. STRAUSS, 1985. *Morphometrics in evolutionary biology.* Academy of Natural Science of Philadelphia. Special Publication 15.

BOOKSTEIN, F. L & R.A. REYMENT, 1989. Microevolution in Miocene *Brizalina* (Foraminifera) studied by canonical variate analysis and analysis of landmarks. *Bulletin of Mathematical Biology,* 51: 657–679.

CAPANNA E., 1981. Caryotype et morphologie cranienne de *Talpa romana* Thomas *de terra tipica. Mammalia,* 45: 71–82.

CAPOLONGO D. & R. PANASCI, 1976. Le talpe dell'Italia centro meridionale. *Rend. Acc. Sci. Mat. Soc. Naz. Sci. Let. Art. Napoli,* 17: 104–138.

CORBET G.B. & J. E. HILL, 1991. *A World List of Mammalian Species.* III ed. Natural History Museum Publications, London, Oxford University Press.

CORTI M. & A. LOY, 1987. Morphometric divergence in Southern European moles Insectivora, Talpidae). *Boll. Zool,* 54: 187– 191.

FILIPPUCCI M.G., G. NASCETTI, E. CAPANNA & L. BULLINI, 1987. Allozyme variation and systematics of european moles of the genus *Talpa* (Mammalia, Insectivora). *J. Mammal.* 68(3): 487–499.

GONZALES J. & J. ROMAN, 1989. Discrimination osteometrica en el genero *Talpa* (Linneo, 1758), en el norte iberico. Doñana, *Acta Vertebrata,* 16(2): 325–327.

HONACKI, J. H., K. E. KINMAN & J. W. KOEPPL, 1982. *Mammal species of the world. A taxonomic and geographic reference.* Allen Press and the Association of Systematic Collections. Lawrence, Kansas, USA. 694 pp.

KRYSTUFEK, B., 1987. Skull variability of *Talpa romana stankovici* from Macedonia. *Acta Theriologia,* 32(28): 463–474.

MARCUS, L. F., 1988. *Automated data acquisition in museums.* Workshop on Computers in Museums, University of Mexico, Spectra.

MARCUS, L. F., 1990. Traditional morphometrics. Pp. 77–122. In: *Proceedings of the Michigan Morphometrics Workshop.* Rohlf F.J. & F.L. Bookstein eds. The University of Michigan Museum of Zoology. Special Publication No.2, Ann Arbor.

MARDIA, K. V. & I. DRYDEN, 1989. The statistical analysis of shape data. *Biometrika,* 76: 271–82.

NEVO, E., 1991. Evolutionary theory and processes of active speciation and adaptive radiation in subternnanean mole rats, *Spalax ehrenbergi* superspecies. Is Israel. *Evol. Biol.* 25: 1– 125.

RAMALHLINHO M. G., 1985. On the taxonomic position of the Portuguese mole. *Arquivos Mus. Bocage* (Ser. A), 3 (1): 1–12.

REYMENT R. A., 1991. *Multidimensional palaeobiology.* Pergamon Press.

ROHLF F. J., 1990. *NTSYS-pc. Numerical Taxonomy and Multivariate Analysis System.* Version 1.60. Exeter Software, New York.

ROHLF F. J. & F. L. BOOKSTEIN (Eds), 1990. *Proceedings of the Michigan Morphometrics Workshop.* The University of Michigan Museum of Zoology Special Publication No. 2: Ann Arbor.

ROHLF, F. J. & D. SLICE, 1990. Extension of the Procrustes method for the optimal superposition of landmarks. *Systematic Zoology,* 39: 40–59.

SAS. 1985. *SAS* Institute, version 6.0. Cary, USA.

TABACHNICK R. E. & BOOKSTEIN F. L., 1990a. Resolving factors of landmark deformation: Miocene *Globorotalia,* DSDP site 593. In: *Proceedings of the Michigan Morphometrics Workshop.* Rohlf F.J. & F.L. Bookstein eds. The University of Michigan Museum of Zoology. Special Publication N.2, Ann Arbor, Michigan.

TABACHNICK R. E., F. L. BOOKSTEIN, 1990b. The structure of individual variation in Miocene *Globorotalia,* DSDP Site 593. *Evolution,* 44: 416–434.

TOSCHI, A. & B. LANZA, 1959. *Fauna d'Italia Mammalia Generalita, Insectivora, Chiroptera.* Calderini:Bologna.

YATES T.L. & D.W. MOORE, 1990. Speciation and evolution in the family Talpidae (Mammalia: Insectivora). In: Nevo E. & O.A. Reig (eds.): *Evolution of subterranean mammals at the organismal and molecular levels.* Wiley–Liss: New York: 1–22.

APPENDIX ONE

KNOWLEDGE FOR NOTHING: ABSORBING INTERNET AND BITNET RESOURCES

JOSE M. BECERRA

Museo Nacional de Ciencias Naturales
José Gutierrez Abascal, 2
Madrid - 28006 (Spain)
MCNJM14@CC.CSIC.ES.

Electronic mail (or "e-mail" as it is known in the media) is a way of communicating among computers located in very different places (even countries!), that allows exchange of messages, documents, programs, and, in modern systems, even graphics.

We are going to discuss here the two major international scholarly networks for e-mail available nowadays: Bitnet and Internet, and the scientific resources that are available there related to the matters treated in this volume.

BITNET (Because It's Time NETwork)

This is the first major network we are going to deal with. This network only allows electronic mail and file transfer, but not in an online fashion. Remote connection to other computers is not possible.

Apart from electronic mail, there is the LISTSERV resource on BITNET.

LISTSERV stands for "list server". Originally, LISTSERV was a mailing-list server which was designed to make group communication easier. People with a common interest were grouped in a list which was then stored on LISTSERV. They could then communicate with each other by sending mail to a special network address. Any piece of mail sent to these special user-ids would then be automatically distributed by the list server to each and every person on the list.

Revised LISTSERV is a brand new list processor which was developed at the Ecole Centrale de Paris in France. It retains the basics of the old LISTSERV and provides good upward compatibility, while offering more sophisticated functions, helpfiles and more user-friendliness.

The usual procedure in order to subscribe to a list is to send email to the userid LISTSERV at the appropriate node. For example, to subscribe to the CONSLINK list, send email to LISTSERV@SIVM, with no Subject line, and in the body of the message put:

SUBSCRIBE CONSLINK Your Name Here

If you are not on a Bitnet node (for example, if you are only connected to Internet), simply append ".bitnet" to the email address, thus: LISTSERV@SIVM.bitnet.

To get a list of forums, you can send the following line in a message aimed to BITNET address NETSERV@BITNIC (if you are in a BITNET node) or NETSERV@BITNIC.BITNET (if you are not) for USA, and NETSERV@EBCESCA1.BITNET for Spain:

SENDME LISTSERV LISTS

You will obtain a group of list addresses that allow you to explore further.

Two interesting documents on Bitnet, that can help you to dive into the electronic networks arena can be obtained from these addresses:

- address: LISTSERV@CMUCCVMA.BITNET
  include in the body of your message the following line:

GET BITNET USERHELP

- address: LISTSERV@BITNIC.BITNET
  include in the body of your message the following line:

GET EMAILNET UPDEGR_D

INTERNET

At the time of its creation (1969), it was called ARPANET (Advanced Research Projects Agency, established by the U.S. Deparment of Defense). Today, Internet comprises a group of networks that provides global access to computing and information resources (Britten, 1990).

In order to have access to Internet resources, you must have a computer (it can be a PC or a Mac) connected to this network, and a resident network protocol (TCP/IP) software, or be logged onto a computer that is an Internet host. The best way to know if you have these requirements is to contact your local computing center.

Once you have an account in your computer center, you will be able to send messages through E-mail from Internet. You will even be able to connect (although just for mail) to Compuserve , a very well-known online service. To send mail from Internet to compuserve, you must use the following format (Schepp, 1990):

CompuServe user ID@compuserve.com

With this connection to the network, you will be dealing (apart from the electronic mail) with two resources present in Internet: TELNET and FTP.

TELNET

It is a remote login protocol in the Internet protocol suite. It allows a user on one host to establish a connection with a remote host and interact as if the user's terminal is connected directly to the remote host. This is also an online facility.

The first thing you need is an Internet address of a TELNET host.

TELNET addresses are of two kinds:

— name.name.name.etc.
for example: boombox.micro.umn.edu
This is called a domain name.

— number.number.number.etc
for example: 128.101.95.95
This is known as an IP address.

To get a list of TELNET hosts, you must send the following line in the body of a message to BITNET address LISTSERV@UNMVM.BITNET:

GET INTERNET LIBRARY (for text file)
or
GET LIBRARY PS (for a postscrip file)

This file contains the St. George, Dr. Art and Mr Ron Larsen, Internet-Accessible Library Catalogs and Databases, 18 pgs, University of New Mexico, Albuquerque and University of Maryland, NM, December 1989.

This guide is an ongoing project listing online library catalogs and databases available within the United States.

Some interesting TELNET places are:

BISON.CC.BUFFALO.EDU	128.205.2.22	SUNY Buffalo Online Catalog (Library)
PAC.CARL.ORG	192.54.81.128	Colorado Assn. of Research Libraries
NYPLGATE.NYPL.ORG	192.94.250.2	New York Public Library (login:NYpl, password: NYpl)
STIS.NSF.GOV	128.150.195.40	NSF's Science and Technology Information System (login: public)
HUB.NNSC.NSF.NET		Wide Area Info Server. Document database Inc. WORLDBOOK, Wall St. Journal (login: WAIS)
EBB.UIT.UNC.EDU		Variety of services (Libtel, WAIS, NetNews)

FTP (File Transfer Protocol)

This service allows a user to transfer files to and from a remote host on the Internet network. You can have access to hundreds of hosts with subjects as diverse as graphics, programming, cartography, music, literature, technical reports, etc. You can compare these services to big stores operating on a self-service approach.

How can you reach this services?

First, we must state that we are only going to deal here with the so-called anonymous FTP; in this kind of FTP, you do not need to have an account in the remote computer. The computer is freely accessible to anyone willing to call it.

All that we have said before about TELNET addresses can be applied here.

***Sources for FTP*:**

There are several ways of finding sources for FTP hosts:

— One of them is to subscribe to a forum list, in which many people share comments, and, from time to time, lists of places for FTP can be obtained. Among them, we can find:

BIOSCI, GIS-L, GRAPHICS, MORPHOMET, etc.

— Another way is to send a message to a BITNET address, asking for a list of FTP hosts.

The address is the following:

LISTSERV@MARIST.BITNET

You must include in the body of your message the following line:

SENDME BITNET FTPLIST

— Another source of information, not only for FTP, but also for many other subjects dealing with networks, is a collection of user introductions called Biobit. This electronic magazine can be obtained through FTP from NIC.FUNET.FI, in the subdirectory pub/sci/molbio/biobit.

— Perhaps the best way to obtain information on FTP places is through Archie.

Archie

Archie is a database that stores information on FTP hosts: addresses, programs, ect. You can make a search on various items (even by country!) in an online session. This database is regularly updated, so you always get the latest information.

So, how do you get access to ARCHIE? If you are Internet connected, it is easy. Telnet to quiche.cs.mcgill.ca (132.206.2.3 or 132.206.51.1, although you must look at the intro screen below, where you will find the most convenient address for you location) and login as user "archie". "help" gets a list of valid commands.

Let us now try a demo session. This example session has been carried out in a PC connected to a VAX, using kermit as communication program[1]. Comments

[1] Default terminal type is usually VT100. Keys many not be all mapped out correctly.

go in brackets and are italicized. My input is in bold. The answer that the machine gives on the screen goes in normal letters.

Beginning of the sample session

PINAR_$ **telnet 132.206.2.3** *[Trying to connect]*
Trying...132.206.2.3
[We are now connected to the remote computer]
Connected to .
Escape character is '^]'.
SunOS UNIX (quiche.CS.McGill.CA)
login: **archie** *[Login to enter into Archie]*
[Now comes an introductory screen to Archie. It is important to read carefully this screen]
ARCHIE: The McGill School of Computer Science Archive Server [2 Apr 1992]

*** Due to a bug, percentages given on the status line may become negative values. This does not affect the search in any way however. We'll put a fix in place later on in the week.
Australian users: archie on archie.au (139.130.4.6), login "archie"
European users: archie on archie.funet.fi (128.214.6.100), login "archie"
UK users: archie on archie.doc.ic.ac.uk (146.169.11.3), login "archie"
Use the 'servers' command to list all archie servers.
A limit of 10 concurrent telnet sessions has been put on archie.mcgill.ca.
Alternative access through the standalone clients available via
anonymous ftp to this machine. See README file in ~archie/clients.

** 'help' for help
** corrections/additions to archie-admin@archie.mcgill.ca
** bug reports, comments etc. to archie-l@archie.mcgill.ca

archie> **help**

Help gives you information about various topics, including all the
commands that are available and how to use them. Telling archie about your terminal type and size (via the "term" variable) and to use the pager (via the "pager" variable) is not necessary to use help, but provides a somewhat nicer interface.
Currently, the available help topics are:

about - a blurb about archie
bugs - known bugs and undesirable features
bye - same as "quit"
email - how to contact the archie email interface
exit - same as "quit"
help - this message
list - list the sites in the archie database

mail - mail output to a user
nopager - *** use 'unset pager' instead
pager - *** use 'set pager' instead
plans - future plans for archie
prog - search the database for a file
quit - exit archie
servers - display a list of all currently available archie servers
set - set a variable
show - display the value of a variable
site - list the files at an archive site
term - *** use 'set term ...' instead
unset - unset a variable
whatis - search for keyword in the software description database

For information on one of these topics type:

help <topic>

A '?' at the help prompt will list the available sub-topics.
Help topics available:

about	bugs	bye	email
list	mail	nopager	pager
plans	prog	regex	servers
set	show	site	term
unset	whatis		

Help topic? **about** *[Let's leave them to explain what they are]*

archie: the McGill School of Computer Science Archive Server Listing Service

Given the number of hosts being used as archive sites nowadays, there can be great difficulty in finding needed software in a distributed environment. You may know that the software that you need is out there, but it can sometimes be difficult to find. The School of Computer Science at McGill University has one solution to the problem - "archie".
archie is a pair of software tools: the first maintains a list of about 1000 Internet ftp archive sites. Each night software executes an anonymous ftp to a subset of these sites and fetches a recursive directory listing of each, which it stores in a database. We hit about 1/30th of the list each time, so each site gets updated about once a month, hopefully balancing timely updates against unnecessary network load.The "raw" listings are stored in compressed form on archie.mcgill.ca (132.206.2.3), where they are made available via anonymous ftp in the directory archie/listings.

The second tool is the interesting one as far as the users are concerned.
It consists of a program running on a dummy user code that allows outsiders to log onto the archive server host to query the database. This is in fact the program we call "archie". Users can ask archie to search for specific name strings. For example, "prog kcl" would find all occurences of the string "kcl" and tell you which hosts have entries with this string, the size of the program, its last modification date and where it can be found on the host along with some other useful information. In this example, you could thus find those archive sites that are storing Kyoto Common Lisp. With one central database for all the archive sites we know about, archie greatly speeds the task of finding a specific program on the net.
Complete anonymous ftp listings of the various sites that we keep in the database may be obtained via the 'site' command and for a list of the sites which we keep track of, see the 'list' command. For a list of all the archie servers worldwide, see the 'servers' command.
archie also maintains a 'Software Description Database' which consists of the names and descriptions of various software packages, documents and datasets that are kept on anonymous ftp archive sites all around the Internet. The 'whatis' command allows you to search this database.
Send comments, bug reports etc to

archie-group@cc.mcgill.ca

If you have a favourite anonymous ftp site that archie doesn't seem to maintain, or if you have additions or corrections to the Software Description database, send mail to

archie-admin@cc.mcgill.ca

archie was written and is maintained by Alan Emtage (bajan@cc.mcgill.ca) and Bill Heelan (wheelan@cs.mcgill.ca). Peter Deutsch (peterd@cc.mcgill.ca) provided (and continues to provide) ideas and inspiration.

Help topic?
archie> **list** *[Now, we shall see how the list command acts. We have cut the listing a bit]*

898 sites are stored in the database

[Domain name]	*[IP address]*	
a.cs.uiuc.edu	128.174.252.1	03:30 20 May 1992
accuvax.nwu.edu	129.105.49.1	03:31 20 May 1992
acsc.com	143.127.0.2	03:31 20 May 1992
adder.maths.su.oz.au	129.78.68.2	03:31 20 May 1992
aelred-3.ie.org	192.48.115.36	03:56 20 May 1992
aeneas.mit.edu	18.71.0.38	03:56 20 May 1992
aerospace.aero.org	130.221.192.10	03:58 20 May 1992
agate.berkeley.edu	128.32.136.1	03:59 20 May 1992

ahkcus.org	192.55.187.25	03:59 20 May 1992
aisun1.ai.uga.edu	128.192.12.9	04:02 20 May 1992
aix.rpi.edu	128.113.26.11	04:04 20 May 1992
aix1.segi.ulg.ac.be	139.165.32.13	05:35 12 May 1992
ajk.tele.fi	131.177.5.20	04:08 20 May 1992
ajpo.sei.cmu.edu	128.237.2.253	04:09 20 May 1992
akiu.gw.tohoku.ac.jp	130.34.8.9	04:14 20 May 1992
alcazar.cd.chalmers.se	129.16.79.30	04:22 20 May 1992
alex.stacken.kth.se	130.237.237.3	04:26 20 May 1992
alf.uib.no	129.177.30.3	04:55 20 May 1992
alfred.ccs.carleton.ca	134.117.1.1	04:57 20 May 1992
algol.cs.umbc.edu	130.85.100.2	04:58 20 May 1992
alice.fmi.uni-passau.de	132.231.10.1	04:59 20 May 1992
.....		
xview.ucdavis.edu	128.120.1.150	03:33 11 May 1992
xylogics.com	132.245.1.95	03:33 11 May 1992
yaouk.anu.edu.au	150.203.4.29	03:33 11 May 1992
zaphod.lanl.gov	128.165.44.202	03:33 11 May 1992
zaphod.ncsa.uiuc.edu	141.142.20.50	03:33 11 May 1992
zariski.harvard.edu	128.103.28.10	03:35 11 May 1992
zebra.cns.udel.edu	128.175.8.11	03:36 11 May 1992
zebra.desy.de	131.169.2.244	18:52 5 May 1992
zeus.cs.umu.se	130.239.32.12	03:36 11 May 1992
zeus.ieee.org	140.98.1.1	03:36 11 May 1992
zug.csmil.umich.edu	141.211.184.2	03:38 11 May 1992

[Now, we are going to search a program that we know has the string "sura" in its file name. We have also cut this list a bit]

archie> **prog sura**

matches / % database searched: 66 /100% *[In less than 2 minutes!]*

[Here comes the first address]
[Domain name, and the IP address in parenthesis]
Host cac.washington.edu (128.95.112.1)
Last updated 13:06 5 May 1992
[What directory is located the file, and information on the accessibility to it (rw–r– –r– –)]
Location: /pub
FILE rw–r– –r– – 863504 Nov 29 22:26 urusei-yatsura.txt
.....
[Here comes the second address]

Host cs.dal.ca (129.173.4.5)
Last updated 13:34 5 May 1992

Location: /pub/comp.archives
FILE r– –r– –r– – 1749 Oct 31 23:18 1991nov1.001543.4799@sura.net
FILE r– –r– –r– – 547 Nov 22 05:18 1991nov22.064744.4716@sura.net
FILE r– –r– –r– – 1081 Oct 17 12:58 1991oct17.152350.6786@sura.net
FILE r– –r– –r– – 934 Oct 19 20:31 1991oct19.224958.19838@sura.net

```
Location: /pub/comp.archives/rec.arts.anime
  FILE    r--r--r-- 3980 Aug 19  1991 ranma-platonic-tsuranuite-translation-lyrics
[Here comes the last address]
.....
Host wuarchive.wustl.edu  (128.252.135.4)
Last updated 04:16 10 May 1992

Location: /mirrors3/rascal.ics.utexas.edu/system7-related
  FILE    rw-rw-r--  6400  Jul 15 1991        Basura_bin
  FILE    rw-rw-r--   231  Jul 27 1991        Basura_intro
  FILE    rw-rw-r--    58  Jul 16 1991        Basura_intro~

archie> quit [We are disconnecting]
Remote connection closed

PINAR_$ [Back again to our local host]
```

End of the sample session

We encourage you to try this database. You will find very valuable information.

FTP sessions

Now you have several addresses of FTP hosts and you are willing to dig into this plethora of program and data sources[2].
We are going to carry out a typical FTP anonymous session. We shall connect to NIS.NSF.NET, where we know of the existence of the document "The Hitchhiker's guide to the Internet", a very valuable report on Internet. The document is stored in the file RFC1118.TXT, in the subdirectory /publications/rfc (we can know this after a search using Archie, or through a message on one list).
As above, the comments will go between square brackets and in italics. Our input goes in boldface, and the output of the remote computer in normal letters.

Beginning of the sampleF FTP session

PINAR_$ *[This is the computer prompt under VMS, the operative system of our computer, a VAX]*

[2] Software is usually kept in compressed formats, which can be identified by a special extension: . zip (for PC), . tar or . Z (for Unix), or. sit (Mac) among others. These files must be downloaded as image files and subsequently uncompressed. Decompression software is usually found in the same places where you FTP, or can be developed as well. Text files should be downloaded as ASCII files.

PINAR_$ **ftp :== ucxftp/ultrix** *[This order allows our computer to emulate Unix-like commands from a VAX. Only necessary for people on machines running VMS!]*

PINAR_$ **ftp nis.nsf.net** *[We are making the call]*
[After a short period of time, the remote host answers our call]
nic.merit.edu FTP server (Version 4.1 Fri Aug 28 11:37:57 GDT 1987) ready.
Connected to NIS.NSF.NET.
Name (nis.nsf.net:mcnjm14): **anonymous** *[The computer asks for our user name; we must type anonymous]*
331 Guest login ok, send ident as password.
Password: *[For password, you can type whatever you want, but it is polite to send your e-mail address]*
230 Guest login ok, access restrictions apply.
[Now, we are connected to the remote computer]
ftp> **dir** *[First, we can look and see what directories there are in the remote computer]*
200 PORT command successful.
150 Opening ASCII mode data connection for /bin/ls.
total 57

[Several UNIX information]			*[size]*	*[file or directory name]*
-rw-r--r--	1 nic	merit	16033 May 19 10:19	$index
-rw-r--r--	1 nic	merit	4870 May 20 14:59	$read.me
drwxr-sr-x	2 nic	merit	512 Mar 16 23:24	acceptable.use.policies
drwxr-sr-x	2 root	system	512 Feb 20 17:02	bin
drwxr-sr-x	3 cise	nsf	512 May 15 19:20	cise
drwxr-sr-x	3 root	system	512 Feb 20 17:02	etc
drwxr-sr-x	7 nic	merit	512 Apr 21 07:26	internet
drwxr-sr-x	2 root	system	512 Feb 20 17:02	lib
drwxr-sr-x	2 nic	merit	512 May 15 17:41	maps
drwxr-sr-x	6 nic	merit	512 Mar 25 08:24	michnet
drwxr-sr-x	11 nic	merit	512 May 13 12:45	nsfnet
drwxr-sr-x	2 omb	omb	512 Apr 28 09:36	omb
drwxr-sr-x	13 nic	merit	512 May 12 22:07	publications
drwxr-sr-x	3 nic	merit	512 Mar 25 08:42	resources
drwxr-sr-x	3 nic	merit	512 Mar 14 19:19	statistics
drwxr-sr-x	3 root	system	512 Feb 20 17:02	usr
drwxr-sr-x	3 nic	merit	512 May 07 16:15	working.groups

226 Transfer complete.
1088 bytes received in 00:00:03.78 seconds
ftp> **cd publications** *[We move to one subdirectory. We know that this is a subdirectory from the information on the list: those entries that in the first column have a 'd' (e.g., drwxr-sr-x), are subdirectories; the rest are files, that you can retrieve]*

250 CWD command successful.

ftp> **dir**
200 PORT command successful.

150 Opening ASCII mode data connection for /bin/ls.
total 111

```
-rw-r--r--  1 nic      merit     2791 May 15 15:28 $index.publications
drwxr-sr-x  2 nic      merit      512 Mar 04 08:31 farnet.gazette
drwxr-sr-x  2 nic      merit      512 May 28 10:42 fyi
drwxr-sr-x  2 iesg     ietf      1024 May 22 11:47 iesg
drwxr-sr-x  2 iesg     ietf     18944 May 22 11:52 ietf
drwxr-sr-x  2 iesg     ietf     12288 May 10 18:43 internet-drafts
drwxr-sr-x  2 nic      merit     2048 May 13 15:37 internet.monthly.report
drwxr-sr-x  2 nic      merit     1024 Apr 03 13:09 linkletter
drwxr-sr-x  2 nic      merit      512 Mar 30 22:02 michnet.news
drwxr-sr-x  2 nic      merit      512 Mar 25 09:24 michnet.tour.guides
drwxr-sr-x  2 nic      merit    14848 May 28 10:04 rfc
drwxr-sr-x  2 nic      merit     1536 May 27 15:22 std
```

226 Transfer complete.
812 bytes received in 00:00:01.58 seconds

ftp> **cd rfc** *[We move down again to the subdirectory we are interested in]*
250 CWD command successful.

ftp> **dir** *[Now, we look for the file we want to retrieve]*
200 PORT command successful.
150 Opening ASCII mode data connection for /bin/ls.
total 78703

[Due to the length of the screen output, we have cut it a bit]

```
-rw-r--r--  2 nic      merit     129670 May 28 10:15 $index.rfc
-rw-r--r--  2 nic      merit       2350 Nov 19 1988  rfc0003.txt
-rw-r--r--  2 nic      merit      26766 Nov 19 1988  rfc0005.txt
-rw-r--r--  2 nic      merit       1585 Nov 19 1988  rfc0006.txt
-rw-r--r--  2 nic      merit       3382 Nov 21 1988  rfc0010.txt
-rw-r--r--  2 nic      merit        367 Nov 18 1988  rfc0016.txt
-rw-r--r--  2 nic      merit       4511 Nov 18 1988  rfc0017.txt
-rw-r--r--  2 nic      merit        310 Nov 19 1988  rfc0018.txt
-rw-r--r--  2 nic      merit       2852 Nov 19 1988  rfc0019.txt
-rw-r--r--  2 nic      merit       2179 Nov 19 1988  rfc0021.txt
-rw-r--r--  2 nic      merit        700 Nov 19 1988  rfc0023.txt
-rw-r--r--  2 nic      merit       3501 Nov 19 1988  rfc0024.txt
-rw-r-r--   2 nic      merit        489 Nov 19 1988  rfc0025.txt
-rw-r--r--  2 nic      merit       3705 Nov 19 1988  rfc0027.txt
-rw-r-r-    2 nic      merit        581 Nov 19 1988  rfc0028.txt
```

.....

```
-rw-r--r--  2 nic      merit      58150 May 19 08:40 rfc1329.txt
-rw-r--r--  2 nic      merit     192925 May 22 10:39 rfc1330.txt
-rw-r--r--  2 nic      merit     129892 May 26 10:10 rfc1331.txt
```

```
-rw-r--r--  2 nic      merit      17613 May 26 10:10 rfc1332.txt
-rw-r--r--  2 nic      merit      29965 May 26 10:10 rfc1333.txt
-rw-r--r--  2 nic      merit      15418 May 26 10:10 rfc1335.txt
-rw-r--r--  4 nic      merit      92119 May 28 10:03 rfc1336.txt
-rw-r--r--  2 nic      merit      22887 May 28 10:03 rfc1337.txt
```

226 Transfer complete.
47562 bytes received in 00:01:09.12 seconds
[Once we have checked that the file we want is there, we proceed to retrieve it]

ftp> **ascii** *[There are two modes for the transmission, depending on the kind of file: ASCII (the default one) for text files, and BINARY or IMAGE for programs and compressed files (see note 2 on page 252). Here, we choose ascii because the file is text. Anyway, you must be aware that for some computers this syntax can change]*

200 Type set to A.
ftp> **get rfc1118.txt** *[This order allows us to transfer the file from the remote host to our local host]*

200 PORT command successful.
150 Opening ASCII mode data connection for rfc1118.txt (61740 bytes).
[Here comes a report on the transaction]
226 Transfer complete.
local: rfc1118.txt remote: rfc1118.txt
63087 bytes received in 00:00:53.74 seconds

ftp> **get $index.rfc** *[It is usually of interest to retrieve a file that contains an index of the files present in the subdirectory, with an explanation of its content[3]]*
200 PORT command successful.
150 Opening ASCII mode data connection for $index.rfc (129670 bytes).
226 Transfer complete.
local: $index.rfc remote: $index.rfc
133153 bytes received in 00:01:55.01 seconds

ftp> **cd /** *[We return now to the root directory]*
250 CWD command successful.

ftp> **bye** *[We disconnect from the remote computer]*
221 Goodbye.

PINAR_$ *[Back again to the prompt of our computer]*

End of the sample FTP session

[3] Other commands:
— *mget* for multiple gets.
— *put* for uploading (given you have permission).

Usually you will have to transfer the files from your local host computer to your PC. For this we use Kermit, a well-known communication program. We have also used Kermit to obtain a log file of the FTP session.

We now give a brief list of places where you can find interesting things for FTP:

Domain name	*IP Address*	*Comments*
ALW.NIH.GOV	128.231.128.251	
WSMR-SIMTEL20.ARMY.MIL		Many interesting programs
FTP.BIO.INDIANA.EDU		
NIC.FUNET.FI	128.214.6.100	
GARBO.UWASA.FI	128.214.87.1	Mirror site of WSMR-SIMTEL20.ARMY.MIL
SDSC.EDU	132.249.20.22	
OAK.OAKLAND.EDU		GraphicsWorkshop (pub/msdos/graphics)
PLAINS.NODAK.EDU	192.33.18.50	ASCII pics, /pub/picture
VAX.FTP.COM	128.127.25.100	FTP software, inc.
ZAMENHOF.CS.RICE.EDU	128.42.1.75	Graphic file format docs.
SBBIOVM.SUNYSB.EDU	129.49.22.2	J.Rohlf morphometric programs (in morphmet.192 directory). Login as GUEST and password ANONYMOU.
HUH.HARVARD.EDU		TAXACOM FTP node

ACKNOWLEDGEMENTS

Thanks to Antonio García-Valdecasas, Les Marcus, and Bill Barnett for their useful comments. Jillian Riordan checked the English. Any 'missaddress' is my own responsibility.

BIBLIOGRAPHY

BRITTEN, William A., 1990. Bitnet and the Internet: Scholarly networks for librarians.*C&RL News*. February: 103-107.

SCHEPP, B., SCHEPP, D., 1990. *The complete guide to CompuServe*. Osborne McGraw-Hill.

APPENDIX TWO

SOFTWARE

LESLIE F. MARCUS

Department of Biology
Queens College of CUNY
Flushing, NY 11367
&
Department of Invertebrates
American Museum of Natural History
CPW at 79th
New York, NY 10024
LAMQC@CUNYVM.BITNET, LAMQC@CUNYVM.CUNY.EDU

The accompanying disk provided, includes programs by Becerra, Marcus, and Rohlf mentioned in various articles. In addition there are updates to some of the programs distributed along with the Proceedings of the Michigan Morphometrics workshop. All of the software is written for IBM PC's or clones.

I. James Rohlf provided the following programs. They are complete programs and are supported by appropriate drivers (in BGIA.EXE) and data files (see the README file along with each program). Programs have been compressed using LHARC.EXE and put in a "self extracting" form so that typing the name of the program will produce the files necessary to run them. To save space, all drivers for monitors and printers have been put in BGIA.EXE. Few of these are needed for any one module depending on your monitor and printer.

Example data files are supplied with each program. It is suggested that each program be put in its own directory on a hard disk, and BGIA.EXE in its own directory. The size of the compressed module and full size of all the components are given below, so they can be run from High Density Floppies as well. Erase the drivers you don't use once you have "extracted" the arced BGIA.EXE and put the appropriate drivers with the programs:

List of Rohlf Programs and Driver File

Self Exploding File	Size	Date Arced	
GRFA.EXE	89490	2-12-93	3:54p
BGIA.EXE	224647	2-12-93	3:58p
TPSRWA.EXE	129914	2-12-93	3:47p
TPSREGRA.EXE	115671	2-12-93	3:56p
TPSA.EXE	115101	2-12-93	3:42p

Note that the A at the end of the file name indicates that they have been compressed into self-extracting files using LHARC.EXE

A list of the files for each program and the sizes are given below:

GRF - Generalized Resistant Fit

GRF.EXE	97952	9-12-91	10:14a
T	3387	10-17-87	11:36a
T1	1697	10-13-88	1:23p
T2	1697	10-13-88	1:28p
T3	3395	10-13-88	1:30p
LIN.DTA	191	7-08-88	4:37p
LIN2.DTA	955	7-08-88	4:37p
LIN3.DTA	955	9-27-88	6:32p
MOSQ.DTA	2705	4-07-88	12:24p
README.GRF	8399	9-12-91	11:26a
GRF.OVR	77765	9-12-91	10:14a

BGI - These are the drivers for graphics monitors and printers

$BMP.BGI	15359	5-28-91	2:19p
$CANON.BGI	17926	5-28-91	2:20p
$CFX.BGI	16876	5-28-91	2:20p
$CGM.BGI	9988	5-28-91	2:21p
$CLQ.BGI	19520	5-28-91	2:21p
$DJ.BGI	17872	5-28-91	2:22p
$FX.BGI	15359	7-26-90	4:48p
$HP7470.BGI	14364	12-27-90	12:01p
$HP7475.BGI	14396	12-27-90	12:01p
$HP7550.BGI	14364	12-27-90	12:01p
$HP7585.BGI	14500	12-27-90	12:02p
$IBMQ.BGI	15359	7-26-90	4:50p
$IMG.BGI	16812	5-28-91	2:27p
$LJ.BGI	16678	10-02-90	11:26a
$LQ.BGI	17358	10-01-90	3:41p
$OKI92.BGI	15359	5-28-91	2:29p
$PCX.BGI	15359	5-28-91	2:29p
$PJET.BGI	18518	7-26-90	4:48p
$PP24.BGI	17214	7-26-90	4:51p
$TIFF.BGI	17272	5-28-91	2:32p
$TSH.BGI	15950	7-26-90	4:48p
$UTIFF.BGI	17272	5-28-91	2:33p
$WPG.BGI	7900	5-28-91	2:33p
ATT.BGI	6348	10-23-90	6:00a
CGA.BGI	6332	10-23-90	6:00a
EGAVGA.BGI	5554	10-23-90	6:00a

HERC.BGI	6204	10-23-90	6:00a
IBM8514.BGI	6665	10-23-90	6:00a
PC3270.BGI	6012	10-23-90	6:00a

TPSRW - Thin Plate Spline Relative Warp

TPSRW.EXE	130752	1-07-93	11:27a
TPSRW.OVR	123652	1-07-93	11:27a
README.TPR	32342	1-07-93	11:24a
RATS7.DTA	3711	5-11-91	10:28p
MOSQ18R.GRF	3468	10-13-92	4:38p
MOSQ18R.LNK	250	10-13-92	4:23p
MOSQ18R.NTS	3359	10-13-92	4:38p

TPSREGR - Thin Plate Spline Regression

This is a new program not discussed in the Valsain or earlier workshops. It is for relating relative warp scores for objects to other characters, or extrinsic variables. See the README file for further details and instructions. This is a very important extension of the spline application software.

RATS.LNK	88	10-12-92	9:54p
RATS.NTS	15809	10-14-92	12:21p
RATS.REF	227	10-14-92	12:24
RATS.SIZ	1734	10-14-92	12:24p
README.REG	18164	12-11-92	12:02p
TPSREGR.EXE	82175	12-11-92	11:39a
TPSREGR.OVR	131952	12-11-92	11:39a
RATS.V1	48	12-11-92	10:39a

TPS - These are the original Thin Plate Spline programs, updated from the Michigan Workshop Proceedings Manual

README.TPS	17307	11-29-91	12:04p
TPSPLINE.EXE	60751	12-17-91	11:26a
TPSPLINE.OVR	83903	12-17-91	11:22a
FIG520.DTA	5255	5-07-89	1:45a
FIG519.DTA	4680	5-07-89	1:45a
FIG518.DTA	5170	5-07-89	1:34a
FIG517.DTA	5196	5-07-89	1:22a
SNEATH3.DTA	4081	5-11-89	2:37p
SNEATH4.DTA	3564	5-13-89	2:00p

SNEATH2.DTA	3929	5-13-89	1:57p
SNEATH1.DTA	3664	5-11-89	2:37p
BKEG1.2	102	10-26-89	2:12a
BKEG1.1	102	10-26-89	12:52a

All above are latest versions of the programs available as this book went to press. Updates are available by anonymous FTP (File Transfer Process) from State University of New York, Stony Brook computer as follows.

1. FTP to SBBIOVM.SUNYSB.EDU
2. Logon as guest... with password..... anonymou
3. change directory to BIOSTAT.192

   cd BIOSTAT.192

4. Look at directory using the DIR command
5. Change the file type mode to binary by typing:

   binary

6. Download the file of interest to your main frame by typing - eg. for the latest TPSRW

   get TPSRWZ.IBMPCEXE

7. Download the TPSRWZ.IBMPCEXE file to your PC or clone. In your PC remember that you are downloading a binary file. For example in Kermit on CMS, you would have to type the special instruction on the main frame SET FILE TYPE BINARY, and make sure that the PC receiving mode is inbinary as well (ask your local computer center for help if in doubt).

8. Rename the program TPSRWZ.ZIP so that PKUNZIP.EXE supplied on the accompanying disk as PKUNZIPA.EXE can "unzip" it. This is best done on a hard disk in a separate directory for each large zipped file.

The Z at the end of the name of the file generally indicates a "zipped" file, as opposed to a self-exploding file as supplied on the accompanying disk. To reiterate, you will have actually a Zipped file and will require PKUNZIP.EXE to decompress it.

II. The second set of programs are MATLAB scripts written by Leslie F. Marcus to accompany his article. In addition there are MATLAB programs for thin pla-

te splines and relative warps which produce the same output as the Rohlf programs. These programs have code which is in the form of the matrix equations in the various articles. However you have to own MATLAB (from Mathworks, Inc.) and have a Math co-processor on your PC or Mac to run them. I have not written graphics for the output - one main point of the "new morphometrics". For research and analysis use the Rohlf programs. To see the actual steps in the analysis the MATLAB programs should be useful.

Programs and relevant data sets are in the self-extracting file:

MATLABA.EXE	15404	2-12-93	3:55p

which includes:

BIPLOT.M	805	7-15-92	11:04a
BIPLOT4.M	2692	7-15-92	10:34a
BIRDLAB.M	131	7-15-92	1:09p
EUEND.M	1193	12-08-92	4:36p
HIPLAB.M	240	7-02-92	5:32p
VARLAB.M	113	5-30-92	2:00p
LHIPPO.DAT	4187	11-15-91	11:03p
LMEDBIRD.DAT	9289	11-15-91	10:38p
LHIPCOL.M	135	6-03-92	11:38p
NUMBERS.M	227	11-25-91	11:12p
MOSQ.M	3177	10-10-92	11:32a
TPLOT.M	454	10-06-92	4:28p
ZYGO.M	4768	9-26-91	4:13p
TPSRAFF3.M	1066	10-10-92	12:15p
TPSRWZ3.M	2259	12-08-92	4:39p
SNEATH1.M	366	12-07-92	10:11p
SNEATH2.M	366	12-07-92	10:11p
SNEATH3.M	366	12-08-92	3:09p
SNEATH4.M	366	12-08-92	3:11p
TPSNEW.M	1945	12-08-92	3:08p
LHIPROW.M	240	6-03-92	11:32p

BIPLOT4.M is the program used to do biplots as in the Marcus article, and BIPLOT.M is a called subroutine. ZYGO.M, LHIPPO.DAT, LMEDBIRD.DAT are the data used in the article. NUMBERS.M, and VARLAB.M are general label files. LHIPROW.M and LHIPCOL.M are row and column labels for LHIPPO. BIRDLAB.M supplies variable labels for the LMEDBIRD data.

TPSRWZ3.M gives similar output to Rohlf's TPSRW program run with MOSQ.M which is a modified MOSQ18R.GRF. MOSQ.M loads the data into an array called "coords". You are asked for the number of specimens, 8, and number of coordinates for each specimen, 18. This corresponds to using a GRF file as input for TPSRW, using the Average as reference, and alpha = 0. TPSRWZ3.M uses subroutines TPLOT.M which plots the reference specimen.

TPSAFF3.M does the affine part of the analysis and is run after TPSRWZ3 as it needs some of the arrays from there. EUEND.M is another data set used in the Talpa article, and must be put in the same format MOSQ.M. The file EUEND.M does just that, so there are several programming steps at the end.

TPSNEW.M gives the same output as Rohlf's TPSPLINE and works with data files SNEATH1.M, SNEATH2.M, SNEATH3.M and SNEATH4.M. These are the same coordinates as in files with TPSPLINE. In each case coordinates are put in an X array in the *.M file.

III. The third program is a useful utility to convert image files as explained in the short text file CONVERIN.TXT found within CONVERTZ.EXE and reproduced between the ..... lines.

.....

CONVERIN.EXE

This program is a utility that works through simple menus. It translates MorphoSys image files to TIFF image files (uncompressed) and viceversa. In this way you can, for example, capture an image with MTV, make a modifi- cation of the image (e.g., histogram flatten), save it, and then input the modified image into MorphoSys. At the moment, the program supports the following resolutions: 512x512, 640x480, and 768x512. The version of MorphoSys version for the European AT-OFG board can not read or write image files.

CONVERTZ when exploded will be found to contain the following files.

CONVERIN.TXT	559	30/01/93	17:38
CONVERIN PAS	13106	10/02/93	8:53
CONVERIN EXE	9152	10/02/93	9:04

CONVERIN.EXE is the executable module and has self contained instructions. CONVERIN.PAS is the Pascal source code for this software. Updates will also be saved and available by FTP from SBBIOVM.SUNYSB.EDU.

OBRAS PUBLICADAS

1. CATALOGO DE LOS FONDOS ESPECIALES DE LA BIBLIOTECA DEL MUSEO NACIONAL DE CIENCIAS NATURALES.

2. LOS REPTILES MESOZOICOS DEL REGISTRO ESPAÑOL.

3. DINAMICA DE POBLACIONES DE PECES EN RIOS.

4. ECOLOGIA DEL OSO PARDO EN ESPAÑA.

5. LA COLECCION DE BRIOZOOS DEL MUSEO NACIONAL DE CIENCIAS NATURALES.

6. CATALOGO DE LOS FONDOS ESPECIALES DE LA BIBLIOTECA DEL MUSEO NACIONAL DE CIENCIAS NATURALES.

7. THE NATURAL HISTORY OF BIOSPELEOLOGY.

8. CONTRIBUTIONS TO MORPHOMETRICS.


CSIC