Semilandmarks in Three Dimensions

Philipp Gunz, Philipp Mitteroecker, and Fred L. Bookstein

T oday there is a fully developed statistical toolkit for data that come as coordinates of named point locations or landmarks. Because all the statistical methods require these landmarks to be homologous among the specimens under investigation it is challenging to include information about the curves and surfaces in-between the landmarks in the analysis. The problem is that these correspond biologically as extended structures rather than lists of distinct points. This chapter is devoted to the method of semilandmarks (Bookstein, 1997), which allows these homologous curves and surfaces to be studied with the existing statistical toolkit. Information from the *interior* of homogeneous tissue blocks is not accessible by these methods.

An earlier morphometric practice uses some nonlandmark points from curves or surfaces as if they were landmarks: the *extremal points* (*Type* III of Bookstein, 1991) that have definitions like "most anterior" or "widest point." These locations, however useful for traditional distance measurements, are ambiguous

Philipp Gunz, Philipp Mitteroecker, and Fred L. Bookstein • Institute for Anthropology, University of Vienna, Althanstrasse 14, A-1091 Vienna, Austria. **Fred L. Bookstein** • Michigan Center for Biological Information, University of Michigan, 3600 Green Court, Ann Arbor, MI 46103.

regarding the one or two coordinates "perpendicular to the ruler." We will call those coordinates *deficient*, and the points to which these coordinates belong, *semilandmarks*. The methodology of semilandmarks this chapter reviews eliminates the confounding influence of the deficient coordinates by computing them solely using the part of the data that is *not* deficient. To be specific, they are treated as missing data and estimated, all at once, in order to minimize the net bending energy (see below) of the data set as a whole around its own Procrustes average. This concept of semilandmarks appeared first in an appendix to the Orange Book (Bookstein, 1991) and was first applied to two-dimensional outline data in Bookstein (1997). Here we explicitly extend the algebra to curves and surfaces in three dimensions and give practical advice on how to collect and interpret this kind of data.

HOMOLOGY

All approaches to landmark-driven morphometrics make one fundamental assumption: that the landmark points are homologous across specimens. The notion of homology invoked in this assumption is not the classic biological notion of that name, which entails similarity of structure, physiology, or development owing to common descent (Ax, 1984; Cain, 1982; Mayr, 1963, 1975; Remane, 1952). In this classic diction, only explicit entities of selection or development can be considered homologous.

Since points *per se* are not likely to be explicit targets of selection, this criterion is too strict—it would rule out almost any use of point coordinates in the course of evo-devo research. Hence for some 30 years morphometrics has used a distinct but related notion of homology, traceable perhaps to an article by Jardine (1969), that centers on variation in the relationships among locations of structures across samples. This notion of homology, often called *geometrical homology*, is embedded in arguments that draw inferences from the appearance of mapping functions, by which we mean the (Cartesian) transformation grid diagrams invented by Albrecht Dürer and rediscovered by D'Arcy Thompson early in the 20th century. The landmarks and semilandmarks that serve as data for the methods of this chapter both arise as careful spatial samples of this underlying mapping function.

For two-dimensional data, landmark locations from photographs or drawings are often sufficient in number to sustain powerful statistical analysis. In three dimensions, however, the number of truly homologous point locations is

often very limited. On the skull, true landmarks are typically located on bony processes, at the intersections of sutures, or at foramina (Richtsmeier et al., 1995). But many curving structures lack punctate landmarks of this sort, and on others candidate points cannot be declared with any assurance to correspond across realistic ranges of variation. The method of semilandmarks begins with structures that are known to correspond as parts (the classic biological notion of homology), and then represents them by geometric curves or surfaces that, in turn, generate reasonable mapping functions. In this way the biological notion of homology has most of its power and sweep restored, as the notion of point-landmark has proved too stringent for effective biometrics in most three-dimensional anthropological applications.

OTHER APPROACHES

There have been earlier attempts to include information from regions lacking landmarks in biometric analysis. Moyers and Bookstein (1979) placed *constructed landmarks* using geometric combinations of defined landmarks along lines erected at specific angles to define new landmarks, but the authors later discarded the method because the prerequisite of homology could not be fulfilled by these new points. Extensions of the thin-plate spline to include curvature information can be found in Bookstein and Green (1993, see also: Bookstein this volume) and Little and Mardia (1996). *Smooth surface analysis* introduced by Court Cutting, David Dean, and Fred Bookstein in 1995 (Cutting et al., 1995) combines the idea of constructed landmarks with previous work on parametric averaging of surfaces (Cutting et al., 1993) for analysis of skull shape in a congenital syndrome, Crouzon Disease. After a thin-plate-spline unwarping to the Procrustes average landmark configuration, equally spaced points are declared homologous along ridge curves and geodesics, and then evenly spaced points are declared homologous on the surface patches lofted above triangles or quadrilaterals woven out of those curves. A statistical analysis separates the total geometric signal into one part from the true landmark points, together with the residual. Andresen et al. (2000) automatically capture semilandmarks using shape features by an algorithm called *geometry constrained diffusion*. Ridge lines, characterized by a minimax property of directional surface curvature, are extracted and matched in order to establish object correspondence. The semilandmarks are mapped into Procrustes space and analyzed using principal coordinates.

Each of these approaches is *ad-hoc* or algebraically inconsistent in one or another important way. There are some Procrustes steps, some Euclidean projection steps, some unwarping steps, and some operations of equal spacing, under the control of no particular governing equation. It would be preferable to have an approach that is matrix-driven at all its steps, so that in studies of modest variation, such as characterizes most quantitative evo-devo work in vertebrate zoology, the variation and covariation of all parameters, whether interpreted, modelled, or discarded as nuisance or noise, can be treated together. To build such a protocol, we exploit the very convenient fact that to the thin-plate spline interpolant, the familiar graphical warping/unwarping operator, there is associated a scalar quantity, the *bending energy*, that is a quadratic form in the locations of the "target" landmark structure. Just as a Procrustes analysis minimizes the sum of squares of a set of forms in one feature space (isometric or affine shape coordinates), so the bending energy can be used to minimize an analogous sum of squares in the complementary feature space of bending, a sum of squares that corresponds surprisingly well to the signals by which features of a geometric homology map are interpreted over a wide range of applications. The combination of these two steps results in an essentially unique set of shape coordinates for the semilandmarks describing most realistic assemblages of landmarks, curves, and surfaces on three-dimensional forms.

WHAT IS WRONG WITH EQUIDISTANT SAMPLES?

To justify a method more complicated than equally spaced points on curves or even triangulations of surfaces, it is necessary to show what goes wrong with those temptingly simple alternatives.

Figure 1a shows a rectangle with one landmark in the lower left corner along with 27 other points spaced equally around the outline. Figures 1b and 1c show a slightly different rectangle with two different sets of semilandmarks. In 1b the points are spaced equally along the outline whereas 1c represents the positions that optimize the bending energy (namely, at zero, for affine transformations). The left thin-plate spline grid in Figure 2 shows a remarkably suggestive pattern of gradients and twists. But since they can *all* be made to disappear by respacing of the semilandmarks on the outline, none of this apparent bending is credible (in the absence of corroborative information, for instance from histology, that some tissue sheet did indeed "turn the corner"). The comparisons we publish, and the statistics that support them, need to apply
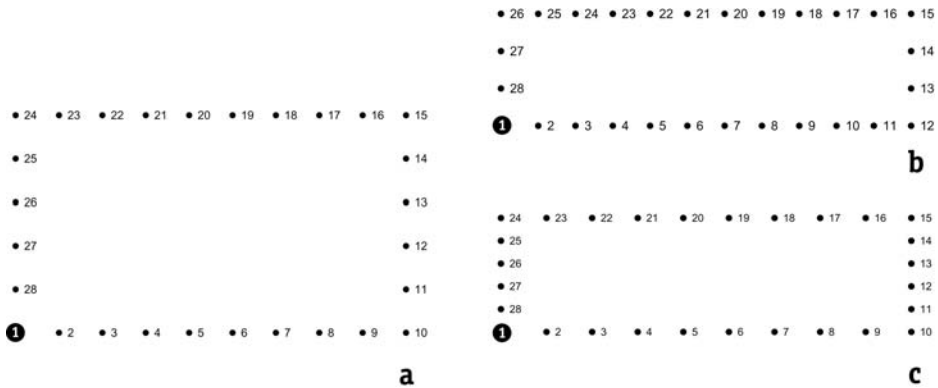
Figure 1. (a) Rectangle with one landmark in the lower left corner along with 27 other points equally spaced around the outline. (b) A more elongated rectangle with the semilandmarks still equally spaced while in (c) the positions are chosen to optimize bending energy (see text).
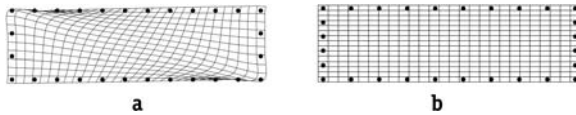

Figure 2. Thin-plate splines corresponding to Figure 1. (a) Deformation grid from the rectangle in 1a to 1b, (b) Deformation grid from the rectangle in 1a to 1c.

in the presence of this ignorance about actual spacing. The only way we can think of to achieve this invariance is to *produce* the spacing as a by-product of the statistical analysis itself.

Figure 3 shows a similar problem for outline structures that bend at large scale. When the points are distributed on the bent form under the criterion of equidistancy (3b), their positions relative to the corners do not correspond to the points in (3a). A better solution is presented in 3c. Figure 4 shows that the TPS grid from 3a to 3c is much smoother (and thus, in this application, less misleading) than the one from 3a to 3b.

While the two generic examples of elongating or bending rectangles might have been resolved in part by placing true landmarks at the corners and *Type III* landmarks at the midpoints of the sides, in many applications Nature is less generous with sharp corners or other shape features that could serve as landmarks. This is the case for the midline of the *corpus callosum*, the structure that connects the two hemispheres of the brain. Figure 5 shows a dataset composed of *corpus callosum* outlines taken from midsagittal sections of MRI scans representing
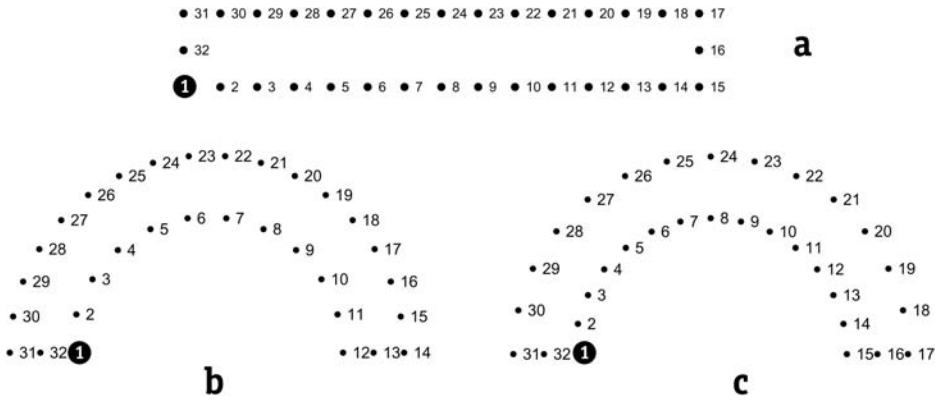
Figure 3. (a) Form with one true landmark in the lower left corner and 31 other points equally spaced along the outline. (b) Bent form with one true landmark (1) and 31 other points in equal spacing. (c) The position of the points now optimizes bending energy.
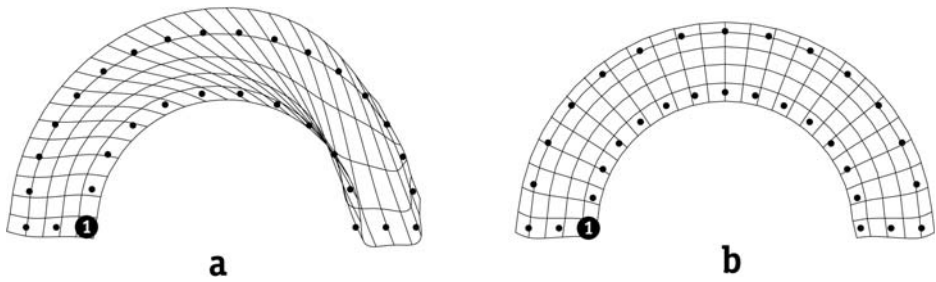

Figure 4. Splines corresponding to Figure 3. (a) Deformation grid from the form in Figures 3a and 3b. (b) Deformation grid from the form in Figures 3a and 3c.
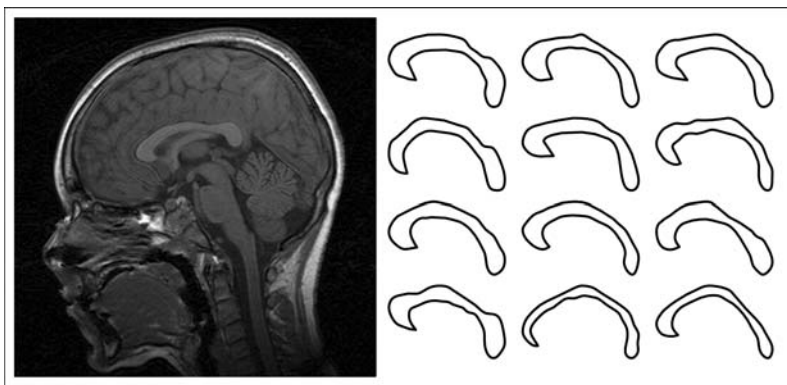

Figure 5. Midsagittal section of an MRI scan and some *corpus callosum* outlines.

normal variation of adults and children. These curves elongate and bend but have only one landmark (*rostrum*).

Figure 6 shows deformation grids between the average (consensus) form and a form with equidistant points compared with the same form captured by semilandmarks. When the consensus is compared to the specimen with the equidistant points, the thin-plate spline deformation grid shows strong local shape differences. Again, there is no reason to consider these changes to be in any way real, as they are very sensitive functions of the arbitrary spacing assumption. By comparison, the points of the form in the lower right corner
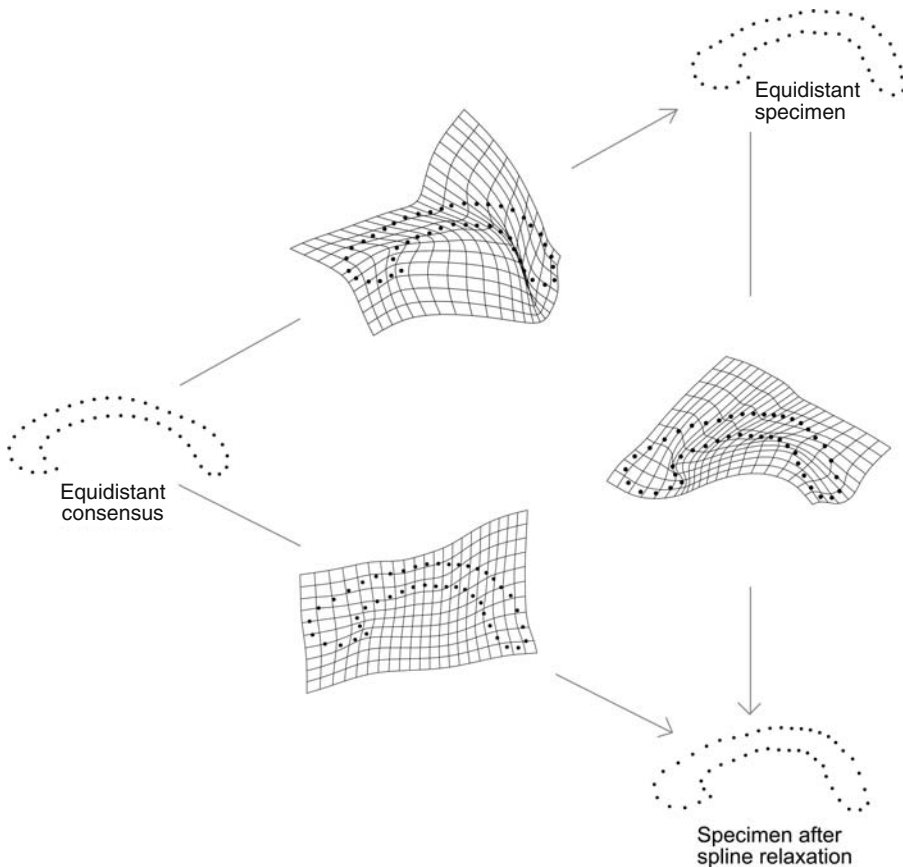

Equidistant
specimen

Equidistant
consensus

Specimen after
spline relaxation

Figure 6. Semilandmarks on the *corpus callosum*. Deformation grids between the consensus form (left side) and a form with equidistant points (upper right) compared with the same form captured by semilandmarks (lower right). Note that the strong local shape effects suggested by the left upper thin-plate spline are an artifact of the equidistancy; the lower left spline, reflecting the real shape difference, is much smoother.

have been placed so as to minimize the bending energy of the interpolation being drawn. Semilandmarks like these can then be treated as homologous, without artifact, in many multivariate analyses, including those that attend to local features of the spline. All those shears along the callosal outline in the upper two grids are meaningless scientifically, regardless of their stark visual effect. These examples typify the ways in which minimizing bending energy serves to protect the scientist from interpreting misleading aspects of a transformation grid in the class of applications concerning us here.

The bending energy that we are minimizing in the course of our analyses is, of course, not itself a biological quantity. It is instead a convenient numeraire for cutting through true ambiguity of empirical representations, rather as the least-squares principle cuts through what would otherwise be the difficult problem of choosing a single line to represent a data scatter. In either case, the aim is to sequester that about which we are truly ignorant (in the linear case, the true errors about predicted values; in the morphometric case, the true spacing of geometric homologues along biologically homologous curves or surfaces). The information that remains stems from the shapes to be studied; arbitrary choices required for digitization have been cancelled out by algorithm. The reason for choosing bending energy instead of, say, Procrustes distance or some other elementary quantity is that in studies where biological interpretation will proceed via features of the grid (rather than, for instance, in terms of phenetic distance or some other narrowly systematic quantity), the bending energy corresponds to the visual signal actually detected by the scientist. It is the local contribution to the variation of second derivatives of the interpolated mapping (see Bookstein, 1991), the rate of change of size and shape of those little grid cells in the deformation diagram, and so is very close to a quantification of the actual information purported to demonstrate any finding claimed. Conversely, bending energy is *invariant* under the operations of a Procrustes superposition—-it doesn't change under rescaling, translation, or rotation of landmark sets—and so computing with it won't interfere with the established Procrustes part of the current geometric morphometrics toolkit.

ALGORITHM

Algebraic Preliminaries

The first two sections following assemble previously published formulas at the core of the method here. This section presents the algebraic setup for the thin-plate spline on landmarks and for the extension to minimizing bending

energy over points sliding on lines, as originally set out by Bookstein (1997). The section on Spline Relaxation on Surfaces shows the notation for the extension to surfaces and section on Flow of Computations sets out the algorithmic cycle we actually follow, which combines these algebraic steps with Procrustes averaging and with projection of semilandmarks from tangent structures back down to actual curving data sets.

In 3D, let U be the function $U(\vec{r}) = |r|$, and consider a reference shape (in practice, a sample Procrustes average) with landmarks $P_i = (x_i, y_i, z_i)$, $i = 1, \ldots, k$. For data in three dimensions, let U be the function $U_{ij} = U(P_i - P_j)$, and build up matrices

$$\mathbf{K} = \begin{pmatrix} 0 & U_{12} & \cdots & U_{1k} \\ U_{21} & 0 & \cdots & U_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ U_{k1} & U_{k2} & \cdots & 0 \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_k & y_k & z_k \end{pmatrix}, \quad (1)$$

$$\mathbf{L} = \begin{pmatrix} \mathbf{K} & \mathbf{Q} \\ \mathbf{Q}^{\mathrm{t}} & \mathbf{O} \end{pmatrix},$$

where $\mathbf{O}$ is a 4×4 matrix of zeros. The thin-plate spline $f(P)$ having heights (values) h_i at points $P_i = (x_i, y_i, z_i)$, $i = 1, \ldots, k$, is the function $f(P) = \sum_{i=1}^{k} w_i U(P - P_i) + a_0 + a_x x + a_y y + a_z z$ where $W = (w_1, \ldots, w_k, a_0, a_x, a_y, a_z)^t = \mathbf{L}^{-1} H$ with $H = (h_1, h_2, \ldots, h_k, 0, 0, 0, 0)^t$. Then we have $f(P_i) = h_i$, all i: f interpolates the heights h_i at the landmarks P_i. Moreover, the function f has minimum *bending energy* of all functions that interpolate the heights h_i in that way: the minimum of $\iiint_{\mathbf{R}^3} \sum \sum_{i,j=1,2,3} \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2$. This integral is proportional to $-W^t H = -H_k^t \mathbf{L_k}^{-1} H_k$, where L_k^{-1}, the *bending energy matrix*, is the $k \times k$ upper left submatrix of $\mathbf{L}^{-1}$, and H_k is the corresponding k-vector of "heights" $(h_1, h_2, \ldots, h_k)$. For morphometric applications, this procedure is applied separately to each Cartesian coordinate: $H = (x_1' \cdots x_k' \, 0 \, 0 \, 0 \, 0)$, then $H = (y_1' \cdots y_k' \, 0 \, 0 \, 0 \, 0)$, then $H = (z_1' \cdots z_k' \, 0 \, 0 \, 0 \, 0)$ of a 'target' form.

In the application to real landmarks, the bending energy of the thin plate spline is the global minimum of the integral squared second derivatives. In the case of semilandmarks this same property can be used as a criterion for optimization: The semilandmarks are allowed to slide along tangents to the curve or surface until the bending energy between a template and a target form is minimal. For curves, we seek the spline of one set of landmarks

$X_1 \ldots X_k$ (the template) onto another set of landmarks $\Upsilon_1 \ldots \Upsilon_k$ of which a subset of m elements are semilandmarks. In the following notation, $i_1 \ldots i_m$ is the list of landmarks that actually slide—this is a sublist of the complete list of landmarks/semilandmarks numbered from 1 through k—so that we use a double notation: Υ_i for the ith landmark/semilandmark, but Υ_{i_j} for the jth sliding landmark. Write Υ^0 for the "starting positions" of all these landmarks. The semilandmarks, Υ_{i_1} through Υ_{i_m}, are free to slide away from their starting positions $\Upsilon_{i_j}^0$ along tangent directions $v_{i_j} = (v_{i_j}^x, v_{i_j}^y, v_{i_j}^z)$ to the curve, while the remaining (nonsliding) landmarks cannot move from their starting locations Υ_i^0. To simplify the following equations, we rearrange the coordinates of all the Υ^0s, sliding or nonsliding, in a vector of the x-coordinates, then the y-coordinates, then the z-coordinates: $\Upsilon^0 = (\Upsilon_1^x, \Upsilon_2^x, \ldots, \Upsilon_k^x, \Upsilon_1^y, \ldots, \Upsilon_k^y, \Upsilon_1^z, \ldots, \Upsilon_k^z)$. To describe the new positions of the m sliding landmarks Υ_{i_1} through Υ_{i_m}, we set out m parameters $T_1 \ldots T_m$ (T for "tangent"), so that the positions after sliding are $\Upsilon_{i_j} = \Upsilon_{i_j}^0 + T_j(v_{i_j}^x, v_{i_j}^y, v_{i_j}^z), j = 1, \ldots, m$. In the ordering of the vector Υ^0, build up a matrix of all these directional constraints together:

$$\mathbf{U}(3k \times m): \quad \begin{aligned} U_{i_j, j} &= v_{i_j}^x \\ U_{k+i_j, j} &= v_{i_j}^y \\ U_{2k+i_j, j} &= v_{i_j}^z, \end{aligned} \qquad (2)$$

where $j = 1, \ldots, m$, all other elements zero.

The sliding now proceeds all at once, all the Υ_{i_j} moving from $\Upsilon_{i_j}^0$ to $\Upsilon_{i_j}^0 + T_j(v_{i_j}^x, v_{i_j}^y, v_{i_j}^z)$, in order to minimize the bending energy of the resulting thin-plate spline transformation as a whole. This bending energy turns out to be

$$-\Upsilon^t \begin{pmatrix} L_k^{-1} & 0 & 0 \\ 0 & L_k^{-1} & 0 \\ 0 & 0 & L_k^{-1} \end{pmatrix} \Upsilon \equiv -\Upsilon^t \mathbf{L_k}^{-1} \Upsilon \qquad (3)$$

in the notation introduced earlier in this section. It has to be minimized over the hyperplane $\Upsilon = \Upsilon^0 + \mathbf{U}T$ and the solution to this weighted least squares problem is

$$T = -(\mathbf{U}^t \mathbf{L_k}^{-1} \mathbf{U})^{-1} \mathbf{U}^t \mathbf{L_k}^{-1} \Upsilon^0. \qquad (4)$$

Anatomical landmarks affect the sliding of semilandmarks, in that they appear in the matrix $\mathbf{L}$ and thus determine the amount of bending energy associated with translations along the tangent vectors T_j semilandmark by semilandmark.

But if you have sufficiently many semilandmarks in general position (at least six points on curves in 2D or 3D, or at least twelve on surfaces in 3D), the semilandmarks can be made to slide "all by themselves," without any need for landmarks to anchor them. For a great deal more explanation of all these matters, the reader is referred to the original journal publication of Bookstein (1997).

Spline Relaxation on Surfaces

The extension of the formalism for surfaces is straightforward: Instead of tangent vectors the semilandmarks are allowed to slide on tangent planes. We seek the spline of one set of landmarks $X_1 \ldots X_k$ (the template) onto another set of landmarks $\Upsilon_1 \ldots \Upsilon_k$ of which a sublist $\Upsilon_{i_1} \ldots \Upsilon_{i_m}$ are free to slide away from their positions along the tangent plane to the surface spanned by two tangent vectors v_{i_j} and w_{i_j} at the original position of the semilandmark. For sliding on tangent planes $\Upsilon_{i_j} = \Upsilon_{i_j}^0 + T_j^1 v_{i_j} + T_j^2 w_{i_j}$, where v_{i_j} and w_{i_j} are unit vectors spanning the tangent plane. Corresponding to the two directions of sliding per semilandmark, the matrix $\mathbf{U}$ of directional information now has two columns per semilandmark: it becomes

$$
\mathbf{U}(3k \times 2m): \quad
\begin{aligned}
U_{i_j,j} &= v_{i_j}^x \\
U_{k+i_j,j} &= v_{i_j}^y \\
U_{2k+i_j,j} &= v_{i_j}^z \\
U_{i_j,j+m} &= w_{i_j}^x \\
U_{k+i_j,j+m} &= w_{i_j}^y \\
U_{2k+i_j,j+m} &= w_{i_j}^z,
\end{aligned}
\tag{5}
$$

where $j = 1, \ldots, m$, all other elements zero.

With this matrix $\mathbf{U}$, equation (4) still supplies the m by 2 matrix of parameters T for which the corresponding semilandmark locations Υ_{i_j} minimize the bending energy (equation 3). Our actual formalism concatenates these two matrices $\mathbf{U}$, one for the curves and one for the surfaces; all the semilandmarks, on curves or on surfaces, slide at once.

Flow of Computations

Our splined semilandmark analysis begins with any convenient selection of semilandmarks on all the curves or surfaces of a data set. The semilandmarks

representing any curve should be equal in number across the sample and should begin in rough geometrical correspondence (e.g., equally spaced); those representing a surface should be reasonably evenly and similarly spaced. Clearly observable curves on surfaces, such as ridges, should be treated as curves instead of surface points; clear local extremes of curvature on curves should be treated as *Type II* landmarks rather than semilandmarks.

The tangents for curves can be calculated as the standardized residual vector of the two neighboring (semi)landmarks. For surfaces the first two principal components of the surrounding landmarks can serve as the two tangent vectors spanning the tangent plane. If the curve or surface is strongly bent in some regions this way of calculating the tangents may become to imprecise. Then either the spacing of semilandmarks should be reduced which results in a larger number of landmarks, or the calculation of tangents should be based on additional information like a denser sampling of curve or surface points or a parametric representation of the curvature (see also section on Data Acquisition).

The *basic algorithm* we propose is then a simple alternation of a Procrustes superimposition with a splined optimization step, each minimizing its own specific sum of squares:

(1) Calculate tangents for each semilandmark.
(2) Relax all specimens against the first specimen.[1]
(3) Compute the Procrustes average configuration.
(4) Calculate new tangents.
(5) Relax all specimens against Procrustes average of step (3).
(6) *Iterate (3)–(5) until convergence.*

During spline relaxation the semilandmarks do not slide exactly on the curves or surfaces but along the curves' or surfaces' tangent structures. Although that reduces the computational effort because the minimization problem is now linear, the sliding along tangents lets the semilandmarks slip off the data. After the relaxation step these points can be placed back on the outline (Figure 7), resulting in a better *extended algorithm*:

(1) Calculate tangents for each semilandmark.
(2) Relax all specimens against the first specimen.

---

[1] There is no initial Procrustes superimposition step necessary because bending energy is invariant to translation, scaling and rotation.

(3) Replace each slid semilandmark by its nearest point on the (curving) surface.

(4) Compute the Procrustes average configuration.

(5) Calculate new tangents.

(6) Relax against Procrustes consensus of step (4).

(7) Replace each slid semilandmark by its nearest point on the surface.

(8) *Iterate steps (4) to (7) until convergence.*

This extended algorithm should be used when sharp curvatures are present in the data set (e.g., the splenium of the *corpus callosum* data set in Figures 5
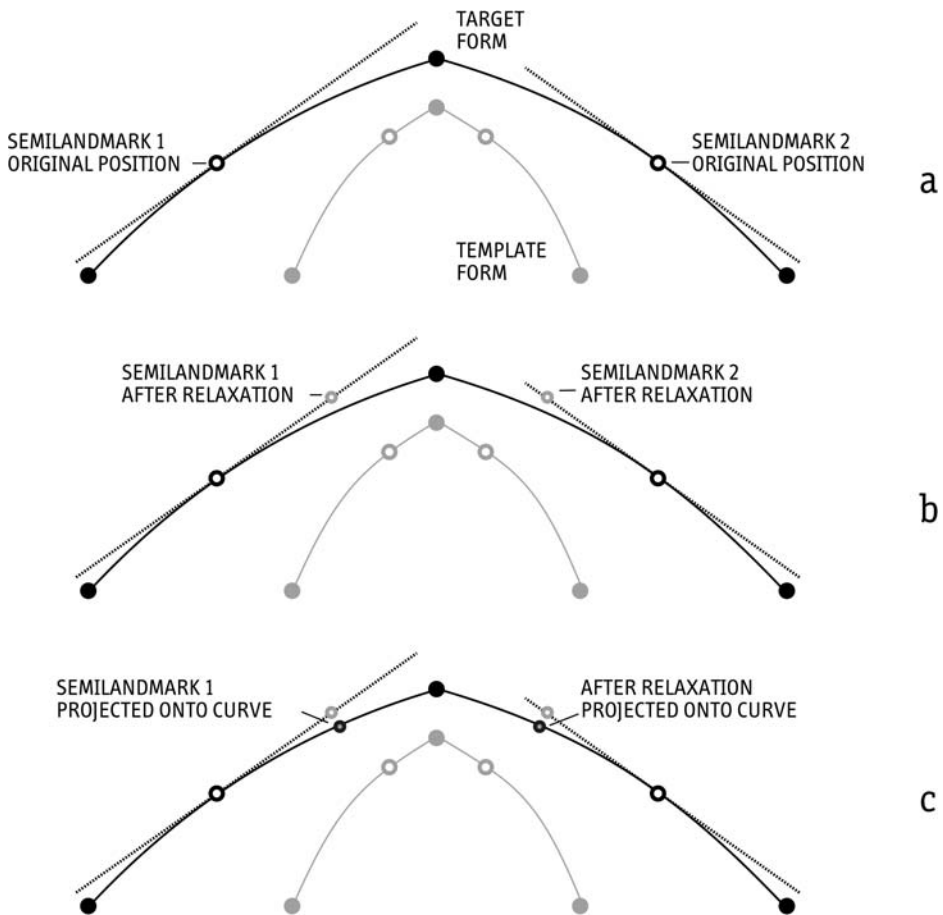

Figure 7.   Sliding along tangents lets the semilandmarks slip off the curve. After the relaxation step these points can be placed back on the outline.

and 6). When applying semilandmarks solely to rather smooth curves or surfaces (e.g., human cranial vault) the basic algorithm usually is sufficient.

STATISTICAL ANALYSIS OF SEMILANDMARK DATA

The more semilandmarks, the better as far as the representation of a geometric form is concerned. In general, the sampling of semilandmarks depend on the complexity of curves or surfaces and the detail of curvature that is of interest. Sampling experiments can help finding an "optimal" number of semilandmarks in the sense of how much information additional landmarks would contribute. For the human neurocranial vault we found 150–200 semilandmarks to be a good representation.

In detailed morphometric data sets, there are far more semilandmarks than specimens (e.g., Bookstein et al., 1999, 2003). This would ordinarily cause a problem for parametric statistical inference, and in the case of semiland-marks there seem to be no actual statistical models available (For instance, Gaussian models for individual semilandmark variation, such as the familiar Mardia-Dryden (1998), do not apply to landmarks bound to lines; notions of independent variation at the multiple semilandmarks of a single curve or sur-face do not apply; etc.). The conventional approach to variable-rich problems, which is to project according to the Procrustes or similarly convenient geometry onto a lower-dimensional empirical eigenspace, will often suffice for such classic comparative themes as allometry or sexual dimorphism. But for more general investigations, it is better to abandon classic statistical models altogether for the more modern alternative that presumes nothing about data distributions at all. Hence excess of variables over cases ends up causing no problems. To pursue this issue (the so-called "high-P low-n issue") would take us far outside the limits of this chapter.

In this model-free context, surveys of empirical data sets proceed by principal coordinates of some distance function (the familiar *relative warps*, for instance, are principal coordinates for Procrustes distance). We don't need to review these methods here (but see e.g., Slice, this volume), as they are the backbone of most of the Procrustes empirical findings ever published; indeed, one principal justification for the semilandmark methods here is that they require no changes whatever in that part of the Procrustes toolkit. Statistical *inference*, on the other hand, requires a somewhat more nuanced adjustment. In our practice, most testing goes via the *randomization methods* first sketched by R. A. Fisher

and now, with the ubiquity of personal computers, perfectly practical for most morphometric studies (Good, 2000).

In general, a permutation test deals with two sets of data vectors for the same specimens. In anthropological applications, one vector will likely be a set of Procrustes shape coordinates for some landmark/semilandmark configuration, and the other vector might be a group i.d. code, another set of shape coordinates, or a collection of non-morphometric measurements. Some statistic relating these two data blocks (such as a group mean difference, or a multiple correlation) is claimed to be interesting and informative, and we want to test this claim against a null hypothesis of no relationship, without making any assumptions whatever about theoretical distributions (Gaussian noise, etc.). We carry out this challenge by considering, or sampling, all the different ways that the rows of the first data matrix could be paired with the rows of the second (i.e., all the permutations of one case order with respect to the other: hence the name of the technique). For each such permutation, compute the same statistic that was claimed interesting in the first place, and collect all the values of that pseudo-statistic (in general there will be $N!$ of them, where N is the total sample size; for a two-group comparison there will be $N!/k!(N-k)!$ nonredundant permutations) in one big histogram. Under the null hypothesis of no meaningful association between the data blocks, the statistic you actually computed should have been drawn randomly from this distribution. So the P-value (technically, the α-level) of the association you actually observed is, exactly, the fraction of this permutation distribution that equals or exceeds the statistic observed. (The word "exact" in the preceding sentence is the same as in the "Fisher exact test" and other familiar contexts. These methods are exact in the sense in which all F-tests and other multivariate Gaussian-assumption approaches are merely approximate under the same conditions.)

For a very small sample with two groups, 3 cases against 3, there are 20 possible rearrangements of the subgroups; thus the best possible P-value you could get is $1/20$, or 0.05. For 8 cases against 8, this minimum P-value is $1/12870$; that is the largest data set for which we have ever computed the exact permutation distribution. For larger samples, the universal custom is to sample from the permutation distribution using a suitable random-number generator (e.g., this is the alternative offered in Rohlf's and Slice's packages available for free at http://life.bio.sunysb.edu/morph/). The observed data set (i.e., the "permutation" with the actual case order preserved between blocks) is to be taken as the first permutation "sampled." For this Monte-Carlo version,

one reports an approximate P-value of m/n with s.e. of $\sqrt{m}/n$, where n is the number of permutations generated and m is the total number of permutations sampled for which the test statistic equals or exceeds the value actually observed. The larger the value of n, the more accurate this approximation.

The power of the test varies by the choice of the test statistic. The authors of this chapter prefer Procrustes distance; others use t-tests, F-ratios, or lower-dimensional multivariate summaries such as T^2. While there is nothing special about a randomization test that is applied to semilandmarks, nevertheless there is something special about the way semilandmarks are used for these statistics. The coordinates that would have been considered "deficient" if these points had been used as landmarks are explicitly *omitted* from statistical manipulations of the resulting Procrustes coordinates. This means, in practice, that the variables consist of distances of the semilandmarks normal to the average curve or surface, or their sums of squares in Procrustes superposition.

WHICH LANDMARKS SHOULD SLIDE?

Bookstein (1991) defined three classes of landmarks, based upon the amount and quality of shape information they represent. Landmarks of TYPE I (juxtapositions of tissues) or II (maxima of curvature) are defined in all coordinates and should as a general rule be taken as real landmarks. TYPE III landmarks, defined by phrases like "the most anterior" or "the farthest from," would better be treated, along with neighboring points, as semilandmarks. Their definitions stem from distance measurements and are therefore informative just in one direction. The other coordinates are deficient and should be estimated by the sliding algorithm whenever using landmark based statistics. Occasionally, however, biological questions warrant the sliding of TYPE II and even TYPE I landmarks that lie on curves and surfaces also captured by semilandmarks. These exceptional landmarks include points on sutures (such as *frontomalare orbitale* on the orbital ridge), particularly crossing points of sutures (such as *lambda* or *bregma* on the neurocranial vault). When the *functional* shape of the neurocranium is of principal interest, landmarks like these should be allowed to slide. Taken as anatomical landmarks, they yield information mainly about development instead (i.e., how the particular neurocranium manages to realize its shape ontogenetically or phylogenetically).

The locations of true landmark points interact with the shape of curves and surfaces in producing the final locations of semilandmarks. Omitting landmarks
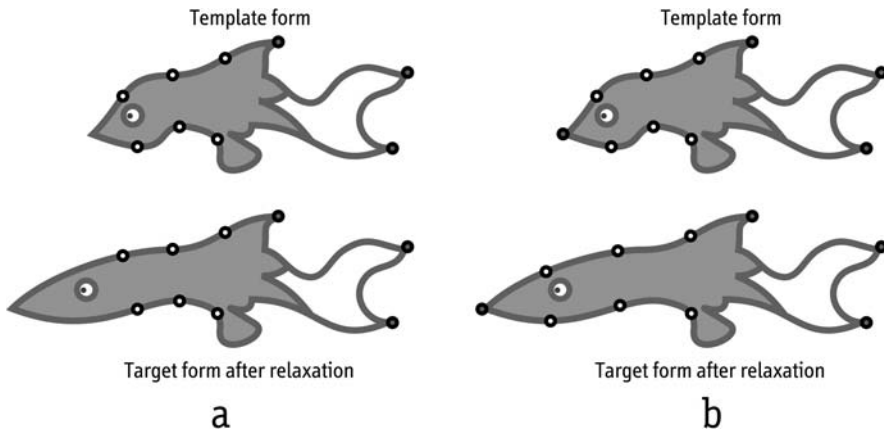
Figure 8. To force a better homology of semilandmarks, use landmarks. (a) For the limited landmark set shown here, minimizing the bending energy slides the semilandmarks (hollow circles) to inappropriate positions. (b) A better set of semilandmarks arises when an additional anatomical landmark (filled circle) is placed at the tip of the "jaw."

when they are easily available, or spacing semilandmarks too sparsely with respect to reliable features of curve or surface form, can produce obviously incorrect results. Figure 8 demonstrates one of these predictable pathologies, as semilandmarks can depart from true landmarks they should accompany or can ignore obvious features of curving form that happen not to have been referred to. We do not set down rules here, as in practice these problems are obvious, once inspected, and the solutions intuitive.

DATA ACQUISITION

The algorithms described above require two kinds of data for each specimen: coordinates of named point locations/landmarks and coordinates of a discrete representation of curving form in-between. In principle there are three types of data sources: discrete landmark point data, discretely sampled curve or surface data, and volume image data. When data begin with image volumes, the first step is usually the explicit location of the curves or surfaces along which semilandmarks will be spaced. This operation is computationally demanding and hardly possible in an algorithmic way, in spite of many experiments in the medical imaging literature. For instance, standard methods for mesh generation fail when patches fold anyway. For a typically *ad-hoc* response to this, see the surface remeshing step in Andresen et al. (2000).

In our experience with primate crania it is faster, more accurate and less expensive to locate samples of semilandmarks explicitly by a device like a Microscribe or Polhemus digitizer that directly yields coordinate data. In particular, one powerful source of information about surfaces is the *ridge curve*, the locus of points with an extreme of surface sectional curvature in the direction perpendicular to the curve. Tracing ridge curves on a virtual specimen is quite tedious, whereas tracing them on a physical specimen is relatively easy. When the physical specimen is not available or one wants to measure internal structures and is hence obliged to use a virtual specimen, we recommend using a software package like Edgewarp3D (Bookstein and Green, 2002) that allows the explicit visualization of sectional curvatures. For surface-semilandmarks one needs to extract a dense cloud of points from the volumetric information—this surface extraction is available in many medical imaging software packages.

How to Measure?

All reasonable approaches to this praxis are constrained by the prerequisites of the semilandmark-algorithm. Semilandmarks have to have the same counts on every curve or surface of the assembly and have to be in the same relative order with respect to each other and to any true landmark points that may be present.

Curves in three dimensions: Procedures for three-dimensional curves are a straightforward extension of those for two-dimensional curves. Although the algebraic formalism does not require the endpoints of curves to be point landmarks, we strongly advise that they be delimited in this way, or else semilandmarks might slip off the available curving data in the course of sliding along tangent lines. To get the same number of semilandmarks in the same order on each specimen, it is convenient to begin with points equidistantly spaced along outline arcs, perhaps through automatic resampling of a polygonal approximation to the curve. In the case of volume image data, one can begin with points spaced inversely to radius of curvature on a typical form, then warp them into the vicinity of every other specimen using only true landmarks, and finally, project them down onto the apparent curve in the image. This is how the curves of Bookstein et al. (2002) were located.

Surfaces: Techniques for surfaces differ substantially from those for curves in that except for planes and cylinders there is no straightforward analogue to the notion of "equal spacing." Along with Andresen et al. (2000), we recommend

beginning with a hugely redundant sample of points on each surface. These can be produced just by "scribbling" around the surface using a device such as a Microscribe digitizer set to stream mode. Alternatively, one can use point clouds generated by a surface scanner or extracted surfaces from volumetric data.

On one single reference specimen, we then carefully produce a mesh of far fewer points, relatively evenly spaced, by thinning the redundant point cloud (Figure 9). (Points should be more dense near ridges of the surface even if those are not to be treated as curves.) The reference specimen is then warped to the landmark configuration of another specimen. On the surface representation of this target specimen the points nearest to the warped mesh are taken as starting positions of the semilandmarks. This procedure is repeated for every specimen in the data set until every specimen possesses a starting configuration for the subsequent relaxation step.

At the same time, the other surface points of each specimen, the ones not used as semilandmarks, continue to supply information for the sliding algorithm; the two dominant eigenvectors of their variation in small neighborhoods around the
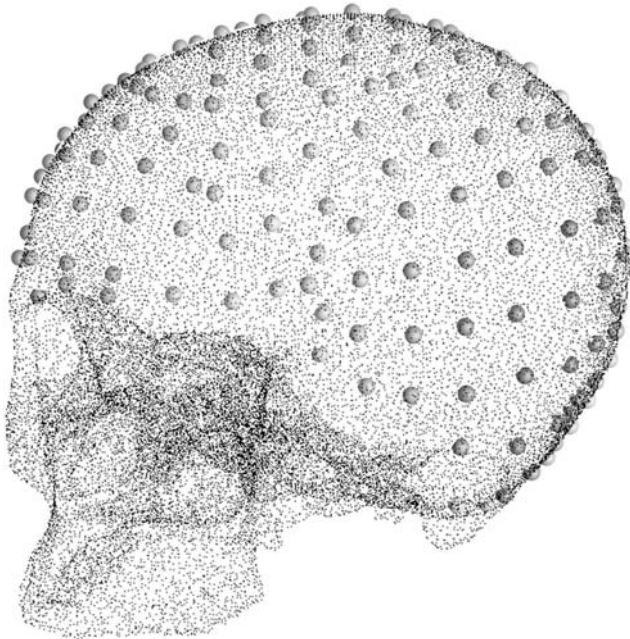

Figure 9. By thinning a dense, discrete representation of the surface we produce a mesh of relatively evenly spaced points, which are then used as starting positions for the semilandmarks (Gunz et al., 2002; Mitteroecker et al., 2004).

semilandmarks are used to specify the vectors v_{i_j}, w_{i_j} of the tangent planes along which they slide. The slid semilandmark can be projected down to the original surface according to the quadric approximation of the surface perpendicular to this best-fitting plane or any other parametric representation like a thin-plate spline.

EXAMPLE

We illustrate the method of semilandmarks using a sample of 52 human crania to study sexual dimorphism (this sample is part of the larger data set of Bernhard, 2003). On each of the 20 adult males, 20 adult females, and 12 subadults we placed 435 landmarks: 37 anatomical landmarks, 162 semilandmarks on three-dimensional-ridge curves, and 236 semilandmarks on surfaces. Most of the anatomical landmarks are in the face and cranial base, with only a few on the neurocranium. The semilandmarks are distributed on seven curves and on the surface of the neurocranial vault. Landmarks and semilandmarks were captured by a Microscribe G2X, and the surfaces resampled as explained in the previous section. All data handling and statistical analysis was done using *mathematica*-routines programmed by the authors.[2] The data set was treated by the basic algorithm described in the section on Flow on Computations. The Procrustes coordinates of the resulting semilandmark locations are shown in Figure 10.

A plot of the first pair of relative warp (RW) scores (Figure 11) shows that the first RW represents ontogenetic development with the children at one extreme and the male adults at the other. Figure 12 visualizes RW1 as a three-dimensional TPS grid computed using all 435 points but drawn as if restricted to the midsagittal plane only. There is general enlargement of the face relative to the neurocranium, marked prognathism, and maxillary extension. Figure 13 visualizes RW2 by the effect of the corresponding TPS on the triangulated surface from one single typical specimen. The effect of RW2 is mostly on relative cranial width.

We performed a Monte Carlo permutation test to assess the statistical significance of the shape difference between adult males and adult females. Using Procrustes distance as test statistic, in 116 out of 3,000 cases the distance between

---

[2] Two-dimensional semilandmarks can conveniently be handled by existing software packages: Bookstein & Green's Edgewarp2D and James Rohlf's "TPS"-programs. Three-dimensional semilandmarks are available in Edgewarp3D (Bookstein and Green, 2002).
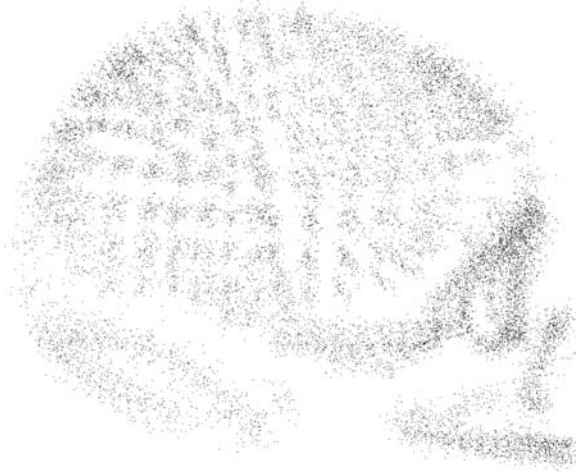
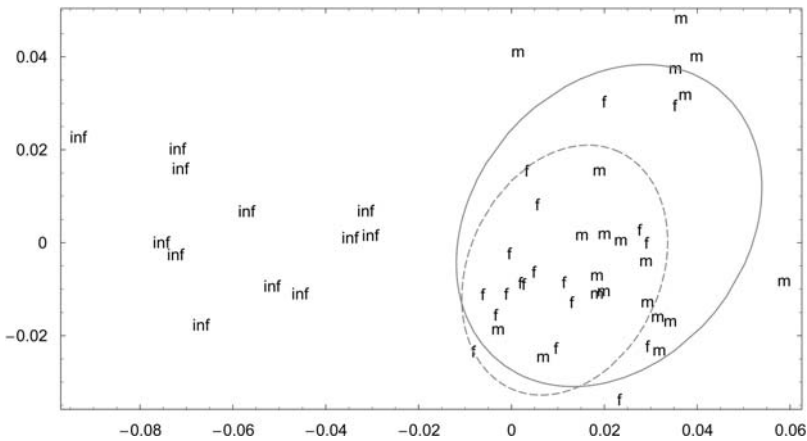**Figure 10.** Procrustes coordinates of 52 *H. sapiens* crania.

Figure 11. Scores of the first relative warp against the second for the full data set. Individuals labelled with "inf" are children, "m" are male adults and "f" are female adults. Note that the 75% confidence ellipse for females (dashed) lies within the male variation (solid ellipse).

randomly relabeled groups was equal to or larger than the actual distance; hence the significance level of the dimorphism is $P \sim 116/3,000 \approx 0.04$. Figure 14 exaggerates this mean difference by a series of factors in both directions. Females have higher orbits and males wider ones; females have a smaller alveolar process, males a broader and more prognathic upper jaw; females have a somewhat globular neurocranial shape, smaller zygomatic arches, and a less pronounced supraorbital torus.
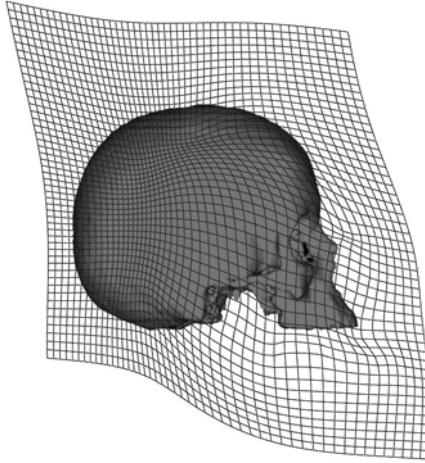
Figure 12. Visualization of the first relative warp (the abscissa of Figure 11) as a midsagittal thin-plate spline. Note the relative enlargement of the face during postnatal development. The specimen shown is the template specimen; the effects of the grid are exaggerated by a factor 4.5.
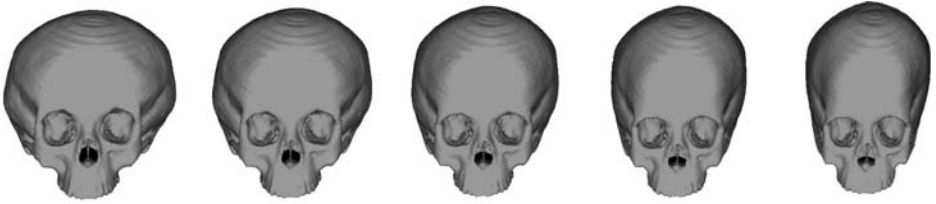

Figure 13. Visualization of the second relative warp as a series of unwarped specimens.

x 7.5 female   x 5   x 2.5   Consensus   x 2.5   x 5   x 7.5 male
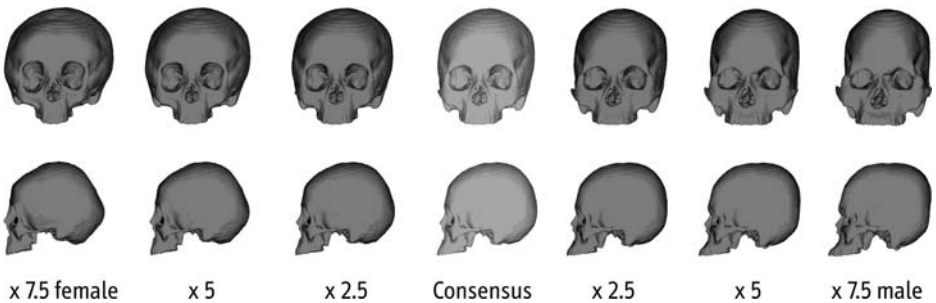
Figure 14. Visualization of sexual dimorphism of *H. sapiens* in a sample of 40 adult specimens. The consensus form in the middle is unwarped to the female mean (left side) and to the male mean (right side). The shape differences are exaggerated to ease interpretation.

P-values like the 0.04 reported just now won't vary much by spacing of evenly spaced semilandmarks, as the Procrustes distances to which they contribute are so redundant. But changes in the coverage of a curving form (addition or deletion of parts, or analysis first by curves and then by surfaces) can alter the strength of statistical findings to an arbitrary extent. There is no general solution to this problem, because a *P*-value is not the answer to any sort of scientific question. As shown, however, the Procrustes methods recommended here result in visualizations of form change in every region of an extended structure. Statistical inferences can go forward quite well in terms of the parts separately even when sliding is in terms of an overall bending energy formalism such as that used here. For an example, see Marcus et al. (1999).

Figure 15 divides this empirical mean difference into a component for static allometry (i.e., the regression of each shape coordinate upon Centroid Size) and a remainder. The upper row shows the relocation of each of the 435 landmarks or semilandmarks that is predicted by sexual size dimorphism; the lower row, the remainder of the actual mean landmark or semilandmark shift between the sexes. The difference between allometry and residual is clearest in the parietal bone,
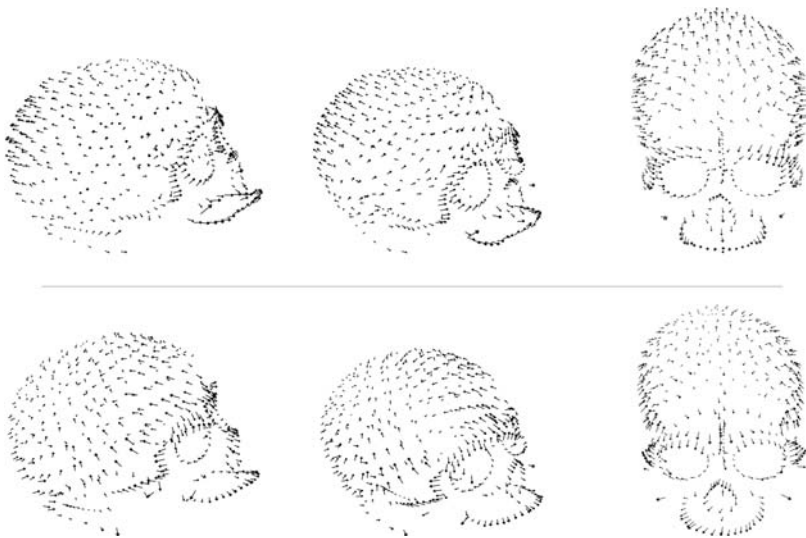

Figure 15.   Sexual dimorphism within the adult subsample, separated into allometric and non-allometric components (see text). The differences between allometry (upper row) and non-allometry (lower row) are most visible in the parietal bone, the zygomatic region, the piriform aperture, and the orbits.

the zygomatic region, the piriform aperture, and the orbits. The multivariate shape vectors for allometry and sexual dimorphism have an angle of 76.3°.

Figure 16 is a different decomposition of the same total sexual dimorphism signal. The left column shows the total mean shift as a little vector at each landmark or semilandmark. When the female consensus configuration is warped to the male using only the true anatomical landmarks, the true landmarks are exactly on the average male position but the semilandmarks' positions are just estimated by the true ones. The middle column of Figure 16 shows this true landmark-driven warping as little vectors—this is the technique of Ponce de Leon and Zollikofer (2001). The picture comes close to the left one because a lot of the information about sexual dimorphism is captured by the traditional landmarks already. The right column shows the residuals from the mean female configuration to the estimated male configuration from the middle column. Notice that many regionally specific aspects of the dimorphism—especially the parietal bosses, the lower temporal bone, the orbits, and the alveolar process—are not accounted for by shifts of landmarks alone. Although a major part of shape change of curves and surfaces in this sample can be reconstructed from landmark positions only, other important local features can be accounted for only by exploiting the additional information in semilandmarks.
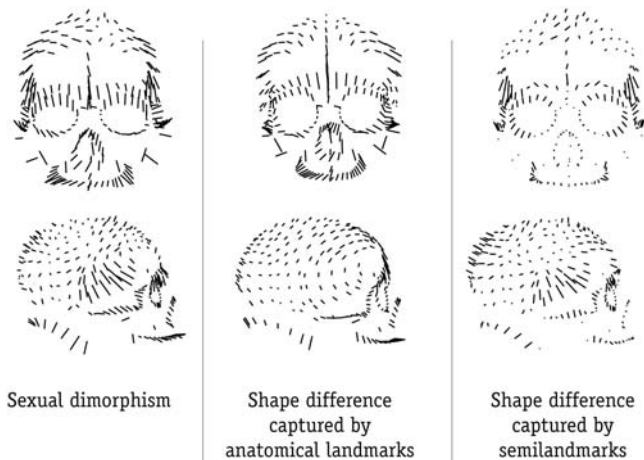

Sexual dimorphism   Shape difference captured by anatomical landmarks   Shape difference captured by semilandmarks

Figure 16.   Sexual dimorphism shown as Procrustes residuals between adult male and female average forms (left). Shape differences that are captured by the anatomical landmarks (middle), and shape differences captured by semilandmarks (right) after a landmark-driven warping. All shifts are exaggerated by a factor 5.

REFERENCES

Andresen, P. R., Bookstein, F. L., Conradsen, K., Ersboll, B. K., Marsh, J. L., and Kreiborg, S., 2000, Surface-bounded growth modeling applied to human mandibles, *IEEE Trans. Med. Imaging* 19:1053–1063.

Ax, P., 1984, *Das phylogenetische System. Systematisierung der lebenden Natur aufgrund ihrer Phylogenese.* G. Fischer, Stuttgart.

Bernhard, M., 2003, Sexual dimorphism in the craniofacial morphology of extant hominoids. Ph.D. Thesis, University of Vienna, Austria.

Bookstein, F. L., 1991, *Morphometric Tools for Landmark Data: Geometry and Biology*, Cambridge University Press, Cambridge (UK), New York.

Bookstein, F. L., 1997, Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Med. Image. Anal.* 1:225–243.

Bookstein, F. L. and Green, W. D. K., 2002, *Users Manual, EWSH3.19*, ftp://brainmap.med.umich.edu/pub/ewsh.3.19.manual.

Bookstein, F. L., Sampson, P. D., Connor, P. D., and Streissguth, A. P., 2002, Midline corpus callosum is a neuroanatomical focus of fetal alcohol damage, *Anat. Rec.* 269:162–174.

Bookstein, F. L., Gunz, P., Mitteroecker, P., Prossinger, H., Schaefer, K., and Seidler, H., 2003, Cranial integration in Homo: Morphometrics of the mid-sagittal plane in ontogeny and evolution, *J. Hum. Evol.* 44:167–187.

Bookstein, F. L., Sampson, P. D., Streissguth, A. P., and Connor, P. D., 2001, Geometric morphometrics of corpus callosum and subcortical structures in the fetal-alcohol-affected brain, *Teratology* 64:4–32.

Bookstein, F. L., Schaefer, K., Prossinger, H., Seidler, H., Fieder, M., Stringer, C., et al. 1999, Comparing frontal cranial profiles in archaic and modern homo by morphometric analysis, *Anat. Rec.* 257:217–224.

Cain, A., 1982, On Homolgy and Convergence, in: *Problems of Phylogenetic Studies*, K. Joysey and A. Friday, eds., Academic Press, London.

Cutting, C., Bookstein, F. L., Haddad, B., Dean, D., and Kim, D., 1998, A spline-based approach for averaging three-dimensional curves and surfaces,

in: *Mathematical Methods in Medical Imaging II*, J. Wilson and D. Wilson, eds., SPIE Proceedings.

Cutting, C., Dean, D., Bookstein, F. L., Haddad, B., Khorramabadi, D., Zonneveld, F. W., and McCarthy, J. G., 1995, A three-dimensional smooth surface analysis of untreated Crouzon's syndrome in the adult, *J. Craniofac. Surg.* 6:444–453.

Dryden, I. L. and Mardia, K. V., 1998, *Statistical Shape Analysis*, John Wiley & Sons, New York.

Gunz, P., Mitteroecker, P., Teschler-Nicola, M., and Seidler, H., 2002, Using Semi-landmarks on surfaces to analyze a Neolithic hydrocephalus, *Am. J. Phys. Anthropol. Suppl.* 33:79.

Mitteroecker, P., Gunz, P., Teschler-Nicola, M., and Weber, G. W., 2004, New morphometric methods in Paleopathology: Shape analysis of a Neolithic Hydroceph-alus, *Computer Applications and Quantitative Methods in Archaeology.*

Good, P., 2000, *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*, Springer, New York.

Jardine, N., 1969, The observational and theoretical components of homology: A study based on the morphology of the dermal skull-roofs of rhipistidian fishes, *Biol. J. Linn. Soc.* 1:327–349.

Marcus, L. F., Frost, S. R., Bookstein, F. L., Reddy, D. P., and Delson, E., 1999, Comparison of landmarks among living and fossil Papio and Theropithecus skulls, with extension of Procrustes methods to ridge curves. Web publication, http://research.amnh.org/nycep/aapa99/aapa6.html.

Mayr, E., 1963, *Animal Species and Evolution*. Harvard University Press, Cambridge, MA.

Mayr, E., 1975, *Grundlagen der zoologischen Systematik*: P. Parey.

Moyers, R. E. and Bookstein, F. L., 1979, The inappropriateness of conventional cephalometrics, *Am. J. Orthod.* 75:599–617.

Ponce de Leon, M. S. and Zollikofer, C. P., 2001, Neanderthal cranial ontogeny and its implications for late hominid diversity, *Nature* 412(6,846):534–538.

Remane, A., 1952, *Die Grundlage des natürlichen Systems, der vergleichenden Anatomie und der Phylogenetik*. Akademische Verlagsgesellschaft Geest und Portig, Leipzig.

Rohlf, F. J., 1993, Relative warp analysis and an example of its application to mosquito wings, in: *Contributions to Morphometrics*, L. F. Marcus, E. Bello, and A. García-Valdecasas, eds., Monografias, Museo Nacional de Ciencies Naturales, Madrid.