Journal of Human Evolution 57 (2009) 48–62

Contents lists available at ScienceDirect

Journal of Human Evolution

journal homepage: www.elsevier.com/locate/jhevol

Principles for the virtual reconstruction of hominin crania

Philipp Gunz [a,b,*], Philipp Mitteroecker [c,d], Simon Neubauer [a,b], Gerhard W. Weber [b], Fred L. Bookstein [b,e]

[a] Department of Human Evolution, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany
[b] Department of Anthropology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria
[c] Department of Theoretical Biology, University of Vienna, Althanstrasse 14, A-1090 Vienna, Austria
[d] Konrad Lorenz Institute for Evolution & Cognition Research, Adolf Lorenz Gasse 2, A-3422 Altenberg, Austria
[e] Department of Statistics, University of Washington, Seattle, WA, USA

ARTICLE INFO

Article history:
Received 21 July 2008
Accepted 12 April 2009

Keywords:
Virtual reconstruction
Geometric morphometrics
Thin-plate spline
Procrustes
Missing data
Semilandmarks
Taung
Nariokotome
Arago

ABSTRACT

Fossils are usually discovered broken or distorted, therefore reconstruction is inevitably the first step towards any comparative analysis. We outline a general methodological framework by which missing information about biological specimens can be estimated using geometric morphometric methods and discuss how this relates to effective paleoanthropological use of incomplete and distorted crania.

Combining digital data resources with geometric morphometrics, we go beyond the assembly of fragments on the computer. As in a three-dimensional jigsaw puzzle, we first assemble the virtual pieces manually. Then we use landmarks, several hundred semilandmarks, and information from complete specimens to estimate missing coordinates and correct for distortion simultaneously. One can thus incorporate information from incomplete specimens in a comparative morphometric analysis while keeping track of the uncertainties that result from partial preservation or deformation. We exemplify our approach by reconstructing the fossil crania Arago XXI, Taung, and KNM-WT 15000. As different assumptions and algorithms lead to different estimations, there exists no "all-purpose" reconstruction. Instead one creates multiple reconstructions—a posterior distribution in a Bayesian sense. This distribution reflects uncertainty due to missing data values and sensitivity to prior assumptions. While there will typically be shape differences among equally plausible reconstructions, these different estimates might still support a single conclusion.

Introduction

Fossils are usually discovered broken or distorted, necessitating reconstruction as an inevitable first step towards any comparative analysis. This is nowadays often accomplished on a computer using virtual specimens acquired with computed tomography (CT), making it possible to work with fossil data without the dangers of harming original materials. "Virtual paleoanthropology" has come a long way since the pioneering work in the late 1970s (Jungers and Minns, 1979; Tate and Cann, 1982; Wind, 1984; Conroy and Vannier, 1984, 1987; Vannier et al., 1985). The first virtual reconstructions of fossil hominid skulls were published by Kalvin et al. (1995). Zollikofer et al. (1995) established software and data management protocols with the same goal and have since continuously refined their methodology (Zollikofer et al., 1998, 2002, 2005; Ponce de León and Zollikofer, 1999; Ponce de León, 2002).

So far, these virtual reconstructions have mostly focused on the "jigsaw puzzle aspect" of assembling fragments on the computer, sometimes using ad hoc methods for correcting post-mortem deformations. Once segmented out of volumetric images, virtual bones are readily assembled manually on the computer, and parts can easily be mirrored and rotated. But those pieces that do not articulate perfectly, either because they are deformed or because parts of their anatomy are missing, can usually be aligned in more than one way. The topic of this paper is what to do *after* the pieces of bone have been assembled based on morphological and anatomical expertise.

If there is more than one way to align fragments, or if parts are missing or deformed, then *any* reconstruction is based on assumptions about functional constraints, integration, symmetry, and sometimes about sex, species affinity, and taphonomy as well. Such assumptions are intrinsic to the reconstruction process, regardless of whether it is done via plaster or via computer.

* Corresponding author.
E-mail addresses: gunz@eva.mpg.de (P. Gunz), philipp.mitteroecker@univie.ac.at (P. Mitteroecker), simon.neubauer@eva.mpg.de (S. Neubauer), gerhard.weber@univie.ac.at (G.W. Weber), flb@stat.washington.edu (F.L. Bookstein).

0047-2484/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jhevol.2009.04.004

Once reconstructed and disseminated via casts, these fossils are often used in analyses without much appreciation of the sometimes considerable uncertainties involved, and with even less appreciation of the sculptor's preconceptions that implicitly shaped the reconstruction. For example, Weidenreich and Swan's widely disseminated composite reconstruction of the Zhoukoudian cranium (Weidenreich, 1937) became the epitome of *Homo erectus* for decades after the loss of the original fossils during the Second World War, even though, as Tattersall and Sawyer (1996:311) dryly noted, it was "largely unconstrained by actual preserved morphologies."

Here we aim to make the reconstruction process transparent and reproducible by formalizing our prior assumptions. We will focus on virtual reconstructions of fossil crania suitable for geometric morphometrics, a set of techniques that has become standard for the statistical analysis of shape and form (Bookstein, 1991; Rohlf, 1993; Bookstein, 1996a; Adams et al., 2004; Slice, 2007). Its analytical tools are built for two- or three-dimensional coordinate data of homologous measurement points called *landmarks*. From an information-theoretic point of view, landmarks that cannot be measured, either because an anatomical structure is broken off or deformed, are missing coordinates. However, the methods and tools commonly used to analyze data matrices with missing values are not readily applicable to morphometric data sets (for reviews and examples from the social and medical sciences see Schafer, 1997; McLachlan and Krishnan, 1997; Schafer and Graham, 2002). Coordinate data measured on biological specimens are quite different from questionnaires and opinion polls, and missing entries in such matrices require special treatment (see Petersen et al., 2006, for a morphometric example). In Gunz et al. (2004) and Gunz (2005) we have laid out a framework for estimating missing values when the data are Procrustes shape coordinates. These methods exploit the fact that the information captured by shape coordinates typically is highly redundant—especially when the measurement points are closely spaced—and take into account prior biological knowledge about functional constraints, curvature smoothness, symmetry, and morphological integration. Here we will focus on the application of these methods to fossil material. Instead of limiting the fossil sample by including only perfectly preserved specimens, one can thus incorporate information from incomplete specimens in the analysis while keeping track of the uncertainties that result from partial preservation or deformation.

Defining and correcting deformation

Parts of an organism that have been deformed due to taphonomic processes cannot be used directly for the analysis, even though the anatomical data are not missing. The following criteria help identify post mortem deformation.

Deviation from symmetry

Crania are approximately symmetric; large-scale deviation from symmetry is a sure indicator of deformation. In this case the diagnostic toolkit provides the remedy at the same time. Restoring symmetry is possible (1) if only one side is affected by distortion and the midline is intact, or (2) if the whole form or parts of it are distorted by a uniform shear.

Biological constraints

In some cases, the morphological constraints imposed upon cranial form by functional requirements, such as mastication, articulation with the vertebral column, and housing of the sensory organs, can be used to identify deformation and then correct for it. If, for example, both maxilla and mandible exist for the same individual, functional integration between the bones, or the wear pattern on corresponding upper and lower teeth could serve both as a guide to identifying deformations and as a criterion for correcting them.

Local distortions of smooth curvature, such as dents and cracks

If CT or surface data are available, then color-coding curvature using a color-gradient can help visually identify such localized distortions. One can approximate a triangulated surface by a quadric form and compute an eigendecomposition. The eigenvalues correspond to the principal curvature values, which will be negative in concave regions and positive in convex regions. Figure 1 shows the surface extracted from a CT scan of the *Homo erectus s.l.* skull KNM-WT 15000 (Walker and Leakey, 1993), with the mean of the principal curvature values colored in Amira.

Exploiting symmetry

The easiest way to deal with missing and deformed anatomical structures is to restore bilateral symmetry. Using digital data sets, it is possible to mirror image the better-preserved side entirely or just parts from one side to the other. If volumetric or surface data are available, then in most applications it will be convenient to restore as much as possible by reflection on the digital specimen, and then subsequently digitize landmark coordinates. Alternatively, it may be possible to digitize 3D coordinates and then reflect these point-locations. There are three ways to reflect points using matrix algebra: (1) mirroring across an empirical midplane, (2) reflected relabeling (RR), and (3) reflection using the thin-plate spline.

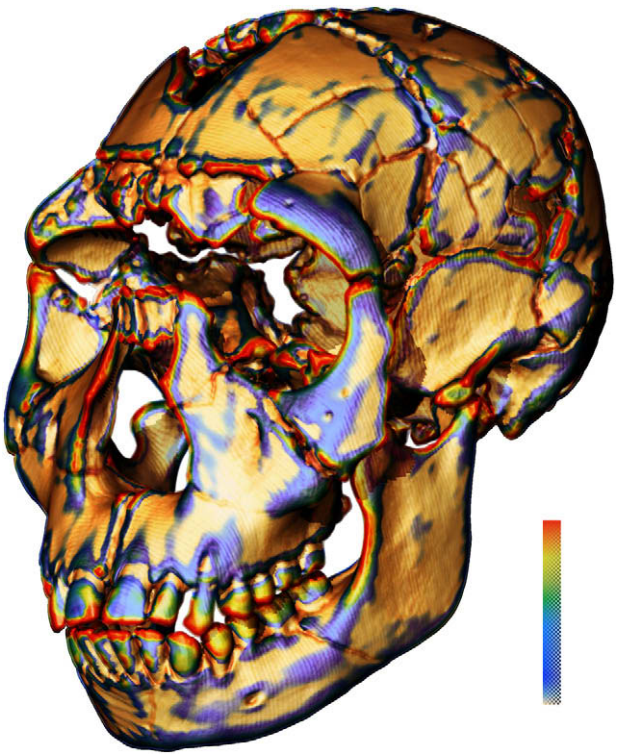

Figure 1. Surface extracted from a CT scan of *Homo erectus s.l.* specimen KNM-WT 15000; the color gradient codes the mean of the two principal curvatures computed for each triangle. Such color-coding can help with visually identifying localized distortions, e.g., breaks and dents.

Mirroring across an empirical midplane

Midsagittal landmarks usually do not lie exactly on a plane, so the symmetry plane has to be estimated (e.g., by a least-squares fit of a plane to the midsagittal landmarks). This can be done by computing the first three principal components of the unpaired landmarks, projecting all landmarks onto the principal component axes (the eigenvectors), and flipping the sign of the scores on the third PC. One then multiplies these flipped points by the matrix transpose of the three eigenvectors calculated previously.

Reflected relabeling (RR)

Rather than defining a specific midsagittal plane, the points can also be mirrored by aligning the landmark configuration with its entire reflection. This requires the specification of paired (bilateral) and unpaired ("midsagittal") points (Mardia et al., 2000; Bookstein and Mardia, 2003; Bookstein, 2005). One reflects the landmarks along any axis while swapping the labels (labels of left and right side landmarks are interchanged); then these two configurations are superimposed with a Procrustes fit based on the available landmarks.

Reflection using the thin-plate spline

Thin-plate spline algebra can be used to deform the reflection of a configuration to fit the preserved anatomy of the original, thereby estimating the unilaterally missing or deformed parts. One computes a thin-plate spline interpolation (Bookstein, 1991) based on the subset of landmarks available on both sides along with the preserved unpaired points. This interpolation function is then used to warp all points of the more complete side onto the other side.

In most practical applications these three methods yield very similar results, yet each has its particular merits. Mirroring across an empirical midplane operates on the assumption that all points on the symmetry axis are close to a single plane and that one side (or part of one side) is better preserved than the other. It is the only available method if only one half is preserved, or if taphonomic distortion is restricted to one side. In the latter case, reflected relabeling and thin-plate spline warping will propagate the error of the distorted side to the other. The biggest advantage of reflected relabeling is that it does not require specifying a mirroring plane. It is very easy to symmetrize the specimen by Procrustes-averaging the form and its relabeled reflection (see example below). If the opposite effect—*preserving* the asymmetry signal—is desired, then reflection using the thin-plate spline is the appropriate method. The reflection may be based on the subset of the "most reliable" (i.e., least distorted) landmarks.

Correcting distortion

Plastic taphonomic distortion affects many fossil specimens and, if left uncorrected, severely limits the choices for comparative analyses (Motani, 1997). As noted above, the effects of plastic distortion can be removed if they are restricted to one side or if the distortion is fairly uniform across the whole specimen. *A priori* biological knowledge is required to pick the "correct" side in the former case, but in practical applications that is rarely a problem. One can then simply mirror the correct side and discard the distorted side. If the distortion is uniform across the specimen, this can be corrected by using reflected relabeling. We will test this approach using simulations in subsequent sections.

Reference-based reconstruction

When the same anatomical parts are missing on both sides or in the symmetry "plane," mirror-imaging no longer suffices to reconstruct the missing data. Instead, one can predict the coordinates of the missing landmarks using information from complete cases. In Gunz (2005) and Gunz et al. (2004) we have presented two algorithmic approaches to complete forms using geometric morphometrics: statistical and geometric reconstruction.

Statistical reconstruction

Statistical reconstruction predicts the location of every missing coordinate using multiple multivariate regressions based on a sample of complete specimens. Each variable with missing values is regressed on all other variables for the complete cases and the missing values are predicted by this linear regression model. Because every coordinate of every landmark and semilandmark is predicted separately, even for large sample sizes, the number of predictor variables will soon exceed the number of specimens, thus making it impossible to invert the variance-covariance matrix. The regressions are therefore performed in a lower-dimensional subspace. To compute this subspace we use the missing and the non-missing landmarks in the reference sample as the two blocks of variables in a two-block partial least squares analysis ("singular warps"; Bookstein et al., 2003). A singular warp is a singular vector from the singular-value decomposition of a covariance matrix between a set of shape coordinates and some other set of variables, possibly another set of shape coordinates (Bookstein et al., 2003; Mitteroecker and Bookstein, 2007; Gunz and Harvati, 2007). The data are projected into a subspace optimized to exploit morphological integration between the available and the missing parts.

Iteratively repeating the regression procedure and incorporating the tentative estimations as predictors resembles the classic expectation maximization (EM) algorithm (Dempster et al., 1977) that is known to yield the maximum likelihood estimate of the missing data under certain distribution models (but see discussion in Gunz et al., submitted for publication).

Geometric reconstruction

Geometric reconstruction is based on the smoothness properties of the thin-plate spline interpolation (Bookstein, 1991). Thin-plate splines are the mathematical basis for deformation grids and semilandmarks (Bookstein, 1997; Gunz et al., 2005) and can be used to "warp" curves and surfaces according to homologous landmarks and semilandmarks from a reference to a target specimen. They can also be used to "unwarp" a volumetric data-set such as a CT scan (Mitteroecker and Gunz, 2009). Semilandmarks are used to quantify curves and surfaces using coordinate-based statistics. Before we detail how this can be exploited for missing-data estimation, it seems appropriate to discuss a few relevant properties of thin-plate splines (TPS). A TPS is computed between two sets of landmark and semilandmark configurations, a reference form and a target form; the "space" between the homologous coordinates is interpolated. Figure 2 shows the difference between warping a surface of a modern human adult to *Australopithecus africanus* specimen Sts 5 based only on 29 anatomical landmarks, and based on 400 semilandmarks along with 29 anatomical landmarks. Despite the considerable shape difference between the modern human reference and the australopithecine target cranium, the match between the two surfaces is almost perfect using semilandmarks. The surface warped according to 29 homologous anatomical landmarks, however, bears little resemblance to Sts 5, especially in the neurocranium. Note that the TPS creates an

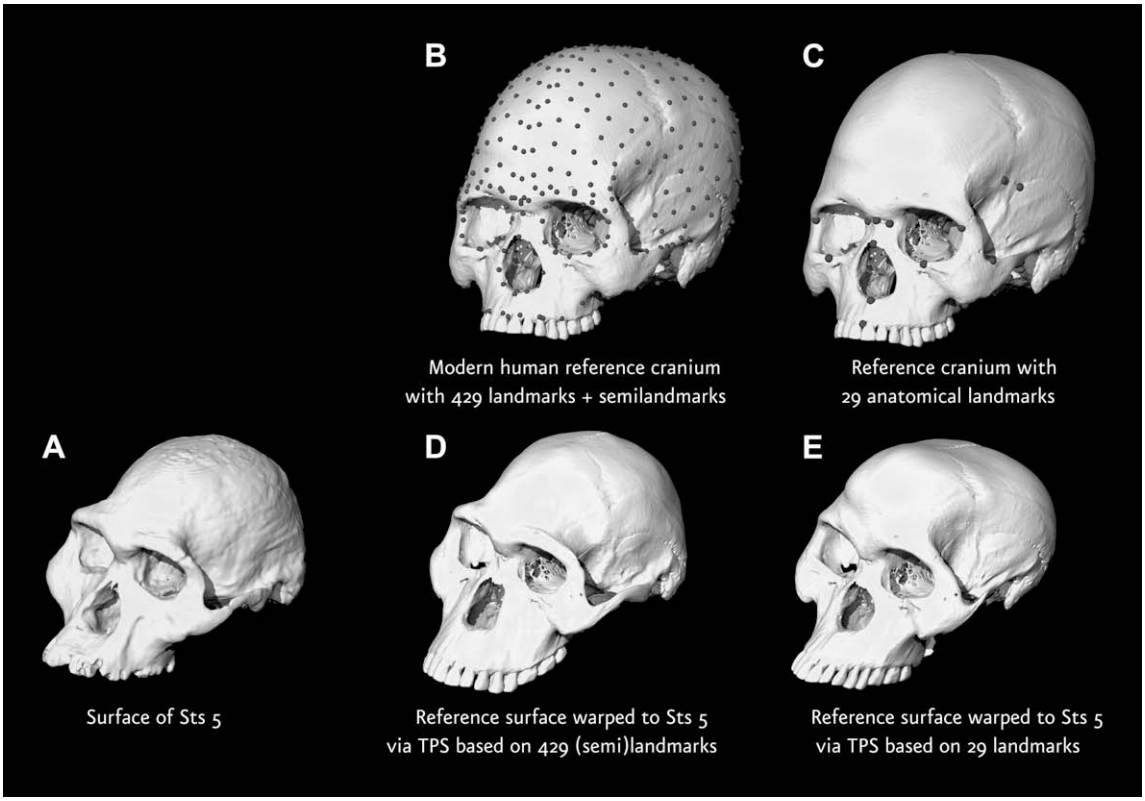

Figure 2. Difference between warping a surface of a modern human adult (B and C) to *Australopithecus africanus* specimen Sts 5 (A) based on 29 anatomical landmarks and 400 semilandmarks (D) and only on 29 landmarks (E). Despite the considerable shape difference between the modern human reference and the australopithecine target cranium, the match between the two surfaces is almost perfect using semilandmarks. The surface warped according to 29 homologous anatomical landmarks bears little resemblance to Sts 5; especially in the neurocranium.

interpolation between only the measured landmarks and semilandmarks; due to the lack of surface-semilandmarks on the maxilla, the warped surfaces in Fig. 2 does not have anterior pillars (Rak, 1983). We will return to the implications for missing data estimation of these important properties of the thin-plate spline in the Discussion.

To estimate missing coordinates on the target form, a thin-plate spline interpolation is computed from the subset of the landmarks and semilandmarks available in both the complete reference and the incomplete target specimen. This interpolation function is used to map the missing landmarks from the reference onto the target.

Algebraically, this can be accomplished by the method of semilandmarks already in use for homologizing curves and surfaces (Bookstein, 1997; Gunz et al., 2005). It requires only a slight modification of the algebra, because one can think of semilandmarks as a method of missing data estimation, based on the TPS. Their position along a curve or surface is not of interest; only their position in the direction normal to the curve or surface is relevant for the statistical analysis. In other words, a semilandmark's position along the curve or surface is *missing*. Therefore the semilandmark algorithm *produces* the point spacing by allowing the points to slide on their respective curve or surface until the thin-plate spline deformation between the specimen and the sample average is least bent. The parameter being optimized is a scalar quantity, the "bending energy" of the thin-plate spline (Bookstein, 1991). A semilandmark on a curve is missing one coordinate, a semilandmark on a surface is missing two. Missing landmarks/semilandmarks are missing all three coordinates. In a TPS-based estimation these missing landmarks/semilandmarks are placed so that the overall bending energy between the reference and the incomplete specimen is smallest, thereby creating a smooth deformation. If a dataset includes semilandmarks, all these missing data can be estimated at once during thin-plate spline relaxation against the reference. The reference configuration can be either a single specimen or a Procrustes group average.

In Fig. 3, we compare the results obtained from statistical and geometric reconstruction using a small dataset of two-dimensional outlines. The data were measured on midsagittal slices of CT scans of 15 adult *Homo sapiens* crania, and comprise a subset of the data used in Bookstein et al. (2003), which should be consulted for details about landmarks and measurement protocol. We deliberately deleted ten semilandmarks on the frontal bone shown as open red circles in Fig. 3C from one of the 15 specimens and reconstructed them (1) by multiple multivariate regression computed using the 14 complete crania, (2) using a TPS based on the non-missing landmarks with the Procrustes mean shape of the complete crania as the reference specimen, and (3) the same but using each of the 14 complete crania as the reference specimen, thereby producing 14 completed forms. Figure 3C and D illustrate that the methods yield very similar results. It is also evident that the choice of the reference specimen changes the reconstruction. The shape variability among the 14 reconstructions in Fig. 3D shows the reconstruction uncertainty that is related to the arbitrary choice of the reference specimen for the TPS in the geometric reconstruction algorithm.

Paleoanthropological applications

Software

We accomplished data processing and analyses in Mathematica, using software routines written by PG and PM. Fossil specimens were measured in Amira and reconstructed in Mathematica and the

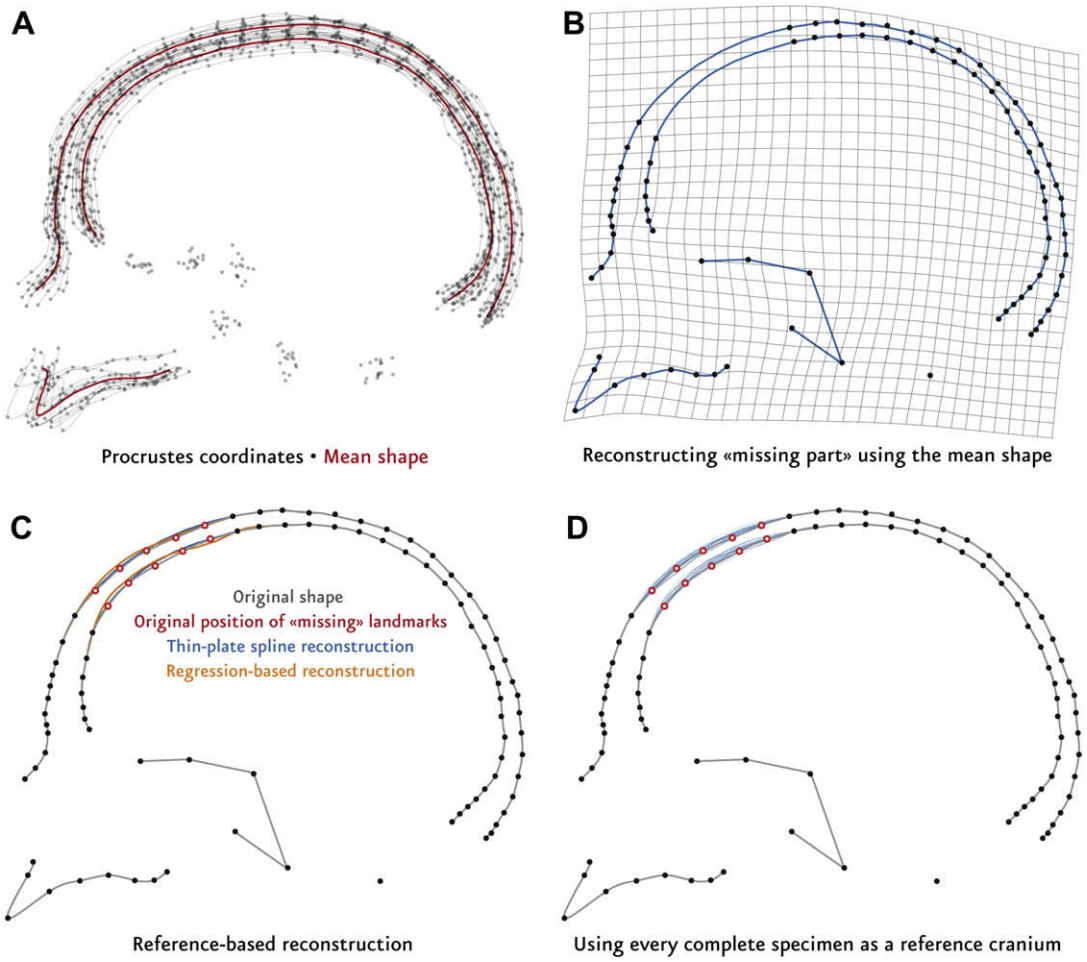

Figure 3. Upper panel: (A) Procrustes coordinates of 15 modern human crania (21 landmarks and 74 semilandmarks) measured on midsagittal sections of CT scans. The Procrustes mean shape is drawn as a red outline. We declared 10 semilandmarks on the frontal bone missing from one individual and estimated the missing coordinates. (B) Thin-plate spline computed from the non-missing parts of the consensus and the "incomplete" specimen. **Bottom panel**: (C) Comparison of the original outline (gray) and the estimates based using geometric reconstruction (blue; warping the Procrustes mean shape using TPS) and statistical reconstruction (orange). The open red circles show the original position of the "missing" coordinates. (D) Using every complete cranium in turn as a reference specimen, we computed 14 estimations of the missing coordinates. This illustrates the reconstruction uncertainty related to the arbitrary choice of a reference specimen in the geometric reconstruction algorithm.

open-source software Edgewarp3D (Bookstein and Green, 2002). Figures were rendered in Mathematica, Amira, and Photoshop.

How much deformation can be corrected reliably?

To test the method of correcting a uniform shear via reflected relabeling (RR), we applied a uniform deformation to the landmarks measured on a CT scan of *Australopithecus africanus* specimen Sts 5 (Broom, 1947). To simulate a uniform shear we multiplied the matrix X of landmark coordinates by

$$Y = X \cdot \begin{bmatrix} 1 & 0 & -0.3 \\ 0.07 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and unwarped the CT volume accordingly via thin-plate spline in Amira (Fig. 4). For details about the difference between warping and unwarping see Mitteroecker and Gunz (2009).

We reflected the distorted landmarks across the x-axis (but reflection across *any* axis would work) and then interchanged the corresponding left and right labels of the bilateral points (we will call this reflected and relabeled configuration Y_{RR}). After a least-squares superimposition of Y and Y_{RR} we computed their Procrustes mean shape, thereby creating a perfectly symmetric form (Mardia et al., 2000). Restoring the centroid size of Sts 5, location and orientation (which were standardized in the course of the Procrustes superimposition) brought this symmetrized configuration into the coordinate space of the original specimen. Finally, we computed a thin-plate spline interpolation between the symmetrized configuration (green in Fig. 4) and the deformed landmark coordinates Y (red in Fig. 4) and used it to unwarp the volume. In Fig. 4C, the original surface of Sts 5 and the corrected surface (green) are superimposed. The corrected surface of Sts 5 is slightly wider than the original; the Procrustes distance between the original and the reconstruction, however, is only slightly larger than the Procrustes distance between the original Sts 5 and its own symmetrized version.

To explore this further, we applied a series of random uniform deformations to Sts 5 in an attempt to find the limits of this correction method via RR. We created 150 increasingly distorted versions of Sts 5 and then corrected them using RR. Figure 5 plots these 150 reconstructions in the Procrustes shape space of 144 crania of *Homo sapiens* (Europeans, $N = 63$; Koe San, $N = 33$), early *Homo* (KNM-ER 1813, KNM-ER 3733, OH 24, KNM-WT 15000, and a reconstruction of "Sinanthropus" described in Tattersall and

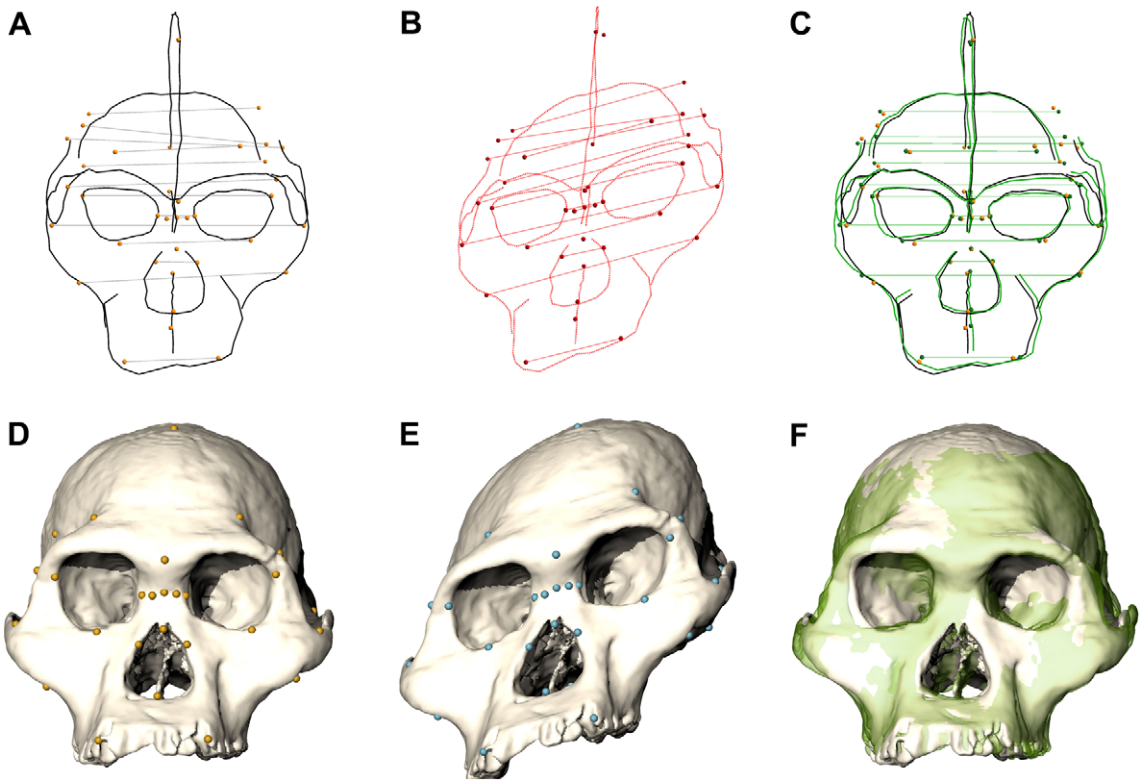

Figure 4. Correcting a uniform shear via reflected relabeling. We measured bilateral and midsagittal landmarks on a CT scan of *Australopithecus africanus* specimen Sts 5 (A) and algebraically simulated a uniform shear (B). Thin lines connect bilaterally symmetric landmarks. C: Original shape vs. the corrected shape after reflected relabeling (green). In the lower panel (D–F) the volume of the original CT scan was unwarped to the deformed and corrected landmark configurations via thin-plate spline interpolation.

Sawyer, 1996), *Pan paniscus* ($N = 36$), and *Australopithecus africanus* (Sts 5, Sts 71, Stw 505). On each cranium we measured 47 anatomical landmarks and 300 semilandmarks (see Gunz, 2005). The 150 corrected specimens were projected into the principal component space computed from the non-deformed crania. Even though the cloud of reconstructions (orange) is quite large, all points fall within the 95% confidence ellipsoid of *Australopithecus africanus*. This pattern is not restricted to the first three principal components: in Fig. 5B the full Procrustes distances between Sts 5 and each of the 150 reconstructions (orange) are plotted together with the Procrustes distances between Sts 5 and all other specimens of the dataset. Only two reconstructions are as far away from Sts 5 as the next closest specimen (*A. africanus* specimen Sts 71). The amount of deformation applied to Sts 5 increases along the x-axis of Fig. 5B; these two "outliers" correspond to shears that were substantially larger than the one shown in Fig. 4.

Unshearing Arago XXI

Even though it is unlikely that real-world taphonomic distortions would result in a perfectly uniform deformation as simulated above, the algorithm can be used to correct crania that display almost uniform shear. Figure 6 shows how the combination of reflected relabeling and TPS warping could be used to remove taphonomic deformation in a fossil cranium. The whole face of Arago XXI (de Lumley and de Lumley, 1971), including the midline, is deformed by what appears to be a uniform shear. A cast of the face of Arago XXI was scanned with a high-resolution optical surface scanner (Breuckmann optoTop-HE) and then digitized in Amira. Note that the cast is of a manual reconstruction that already corrects a local plastic deformation on the frontal bone of the original specimen.

We measured points along the midsagittal curve along with bilateral landmarks and symmetrized them by RR. Then we computed a thin-plate spline interpolation between the original landmark coordinates and the symmetrized configuration and used it to warp the vertices of the triangulated surface. Because there exists one-to-one correspondence between the vertices before and after warping we could even use the texture of the original surface scan.

Reconstructing Taung via TPS

The Taung skull (Dart, 1925), the type specimen of *Australopithecus africanus*, consists of a facial, a mandibular, and a neurocranial segment. The latter comprises a partially preserved natural endocast, but most of the exterior neurocranial shell is missing. In order to include this cranium in a larger dataset of extant and extinct primates (Gunz, 2005), we reconstructed the missing parts of the exterior bony shell of the braincase. The original fossil was scanned at the Sunninghill Hospital, Johannesburg, South Africa (the digital data can be obtained from www.virtual-anthropology.com). The mandibular fragment was not used in this reconstruction. Using Amira, the CT scan of the natural endocast was manually translated and rotated so as to exactly align the fractured surfaces of the two pieces of the endocast. The frontal part of the natural endocast, which is still embedded in the facial fragment, was then extracted by manually segmenting every slice of the CT scan. For further analysis, triangulated surfaces were extracted from the resulting CT images.

In Amira, we then used mirror-imaging across a best-fit plane to the midsagittal landmarks to symmetrize the cranium, thereby completing the left side (Fig. 7A). Along with some taphonomic deformation, this reconstruction thus also discards the biological

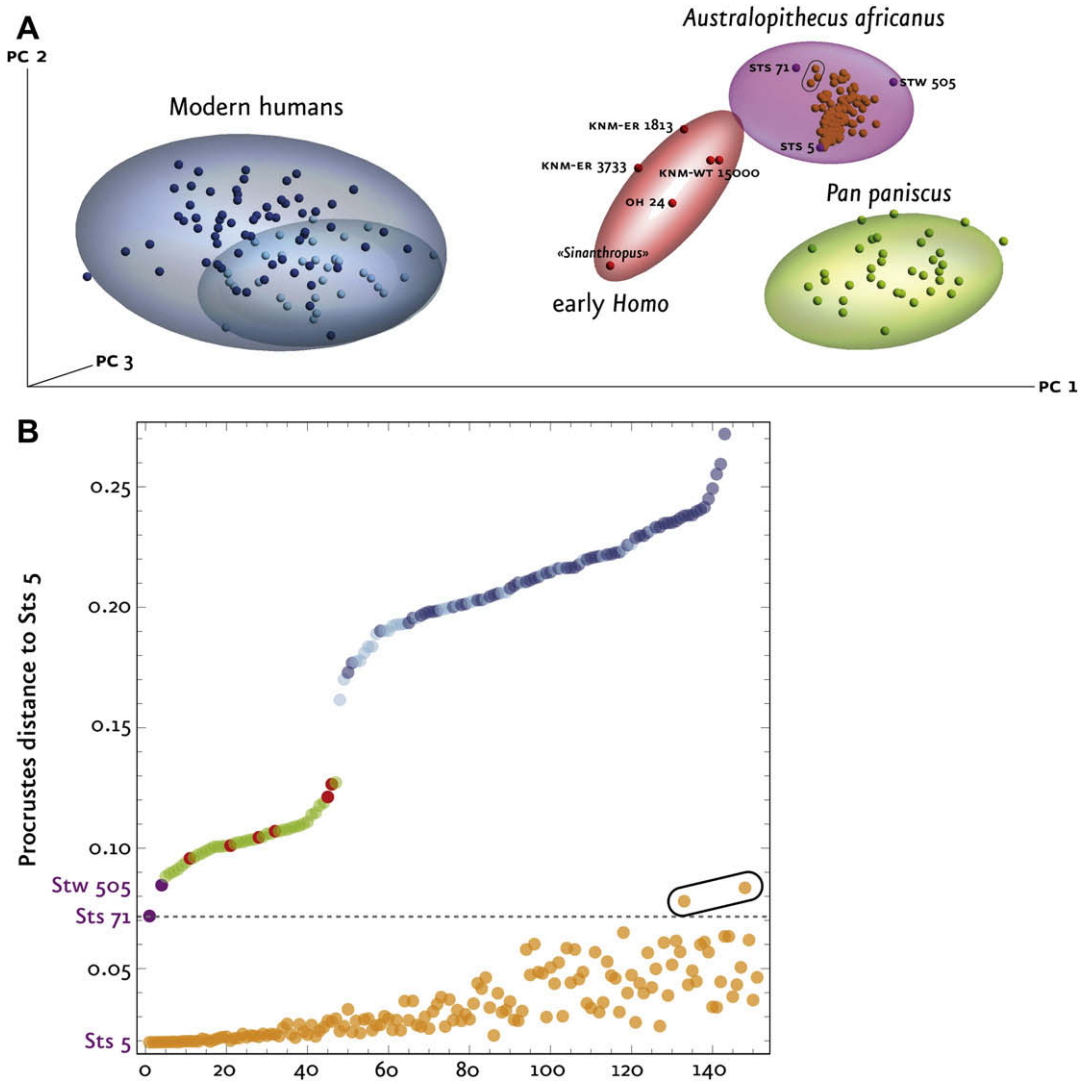

Figure 5. We created 150 increasingly sheared versions of Sts 5 and corrected them using reflected relabeling. A: Cloud of 150 corrected configurations of Sts 5 (orange) projected into the first three principal components of Procrustes shape space of modern humans (blue), bonobos (green), *A. africanus* (purple), and early *Homo* (red); 95% confidence ellipsoids are plotted in the same color as the groups' points. Note that the more heavily distorted specimens plot farther away from the original Sts 5 but all fall within the 95% confidence ellipsoid of *A. africanus*. B: (Sorted) Procrustes distances between all specimens and Sts 5 (y-axis), again colored by group. The corrected versions of Sts 5 are plotted in orange; amount of deformation increases along the x-axis. Only two reconstructions are closer to another specimen (Sts 71) than to Sts 5.

asymmetry of the endocast. Anatomical landmarks ($N = 25$; Table 1) were then measured on this symmetrized version of the Taung child's surface and on a complete reference cranium, a modern human child at the same dental stage. We further digitized curve semilandmarks along ridge curves (Table 2) of *both* specimens as well as surface semilandmarks on the interior and exterior surfaces of the calvaria of the reference specimen. In the subsequent reconstruction steps, the thin-plate spline interpolant was used as the criterion for geometric correspondence: points between the two specimens were considered geometrically corresponding when the thin-plate spline between the reference specimen and the resulting configuration was least bent in the sense of having minimum bending energy.

In the software package Edgewarp3D, we first computed a thin-plate spline interpolation based on the 25 anatomical landmarks to warp all semilandmarks and missing landmarks from the reference specimen into the space of the Taung cranium. This was accomplished by declaring all semilandmarks missing or "fully relaxed." At this stage the anatomical landmarks match perfectly, while all semilandmarks are crudely estimated according to minimum bending energy of just the landmarks, which means that they "float" in the vicinity of their respective curves and surfaces. We then projected each curve-semilandmark onto the preserved part of its respective curve and each surface-semilandmark onto the preserved part of its respective endocranial or ectocranial surface in the Taung child. Semilandmarks that would cover the missing anatomy (i.e., essentially the complete exterior surface of the parietal and occipital bones) were manually declared missing. Then we allowed the semilandmarks to slide along tangents to the curves and surfaces, whereas all points coded as missing were "fully relaxed." After sliding, semilandmarks were projected back onto the preserved curve or surface. Because we used the landmarks and semilandmarks to warp the triangle vertices of the reference specimen's surface, declaring a coordinate missing or present resulted in a continuous update of the reconstructed form. We repeated the spline relaxation and projection steps until convergence, determining by inspection whether or not our semilandmark coding as "missing" or "present" made anatomical sense.

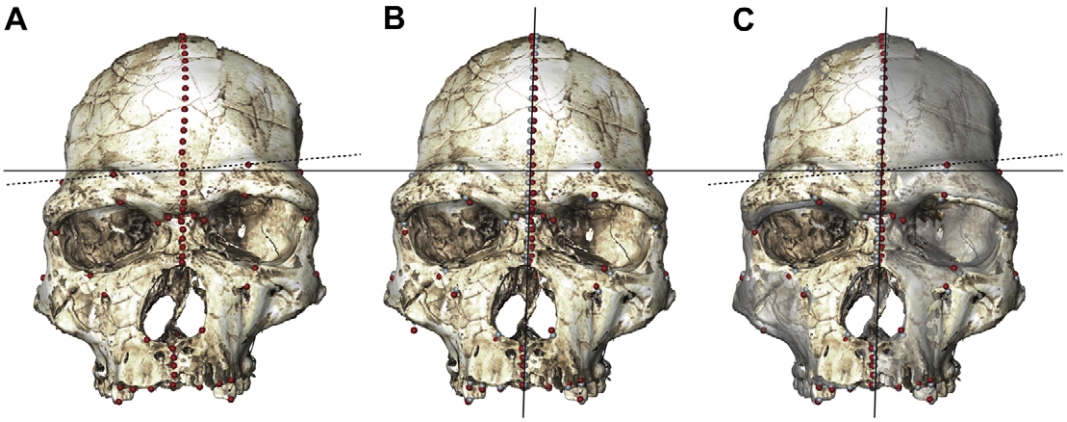

Figure 6. (A) The face of Arago XXI with bilateral and midline points. (B) The points are symmetrized by superimposing them with their relabeled reflection and then computing an average form (blue landmarks). (C) The surfaces before (grey transparent) and after (with texture) correcting the distortion are superimposed to highlight the differences. The advantage of this technique over conventional mirror-imaging is that it can handle deformed "midplanes." The dashed and solid lines are merely illustrations, not part of the computation.

Figure 7 shows the point set used for reconstruction; anatomical landmarks are drawn in blue, curve semilandmarks in gray, surface semilandmarks in orange, and missing points in red. The final version of the reconstruction is drawn as a transparent surface (Figs. 7 and 8).

Naturally, the choice of the reference specimen influences the shape of the reconstruction. We repeated the reconstruction protocol of the Taung cranium using a juvenile chimpanzee as the reference cranium; the resulting forms are very similar, as shown in Fig. 8. The shape differences between the two reconstructions are very subtle, and are visible only in the mastoid region, where they reflect the differences between the reference crania. These shape differences may not matter for a particular analysis when all reconstructed shapes support the same conclusion.

Reconstructing the Turkana boy

One can take the idea of using more than one reference cranium one step further and use *all* possible reference crania in the dataset, thereby creating a distribution of reconstructions. We demonstrate this approach with a fossil reconstruction of a *Homo erectus s.l.* skull from Kenya, the "Turkana boy" KNM-WT 15000 (Walker and Leakey, 1993). The skull of this fossil specimen is almost complete, but the supraorbital torus is missing, and the right orbital margin is damaged.

On a high-resolution CT scan of the *Homo erectus s.l.* fossil we corrected the deformation of the right orbit by mirror-imaging the undistorted left side; this was done in the software package Amira. We then measured a set of 47 anatomical landmarks and 300 semilandmarks on curves and the neurocranial surface (Fig. 9). The same landmarks and semilandmarks were measured on 250 complete crania comprising *Pan troglodytes* ($N = 67$), *Pan paniscus* ($N = 36$), gorillas ($N = 45$), two modern human populations (Europeans, $N = 63$; Koe San, $N = 33$), and 5 fossil crania of early *Homo*. The recent crania were adults of either sex without missing data. See Gunz (2005) for details about the sample and measurement protocol.

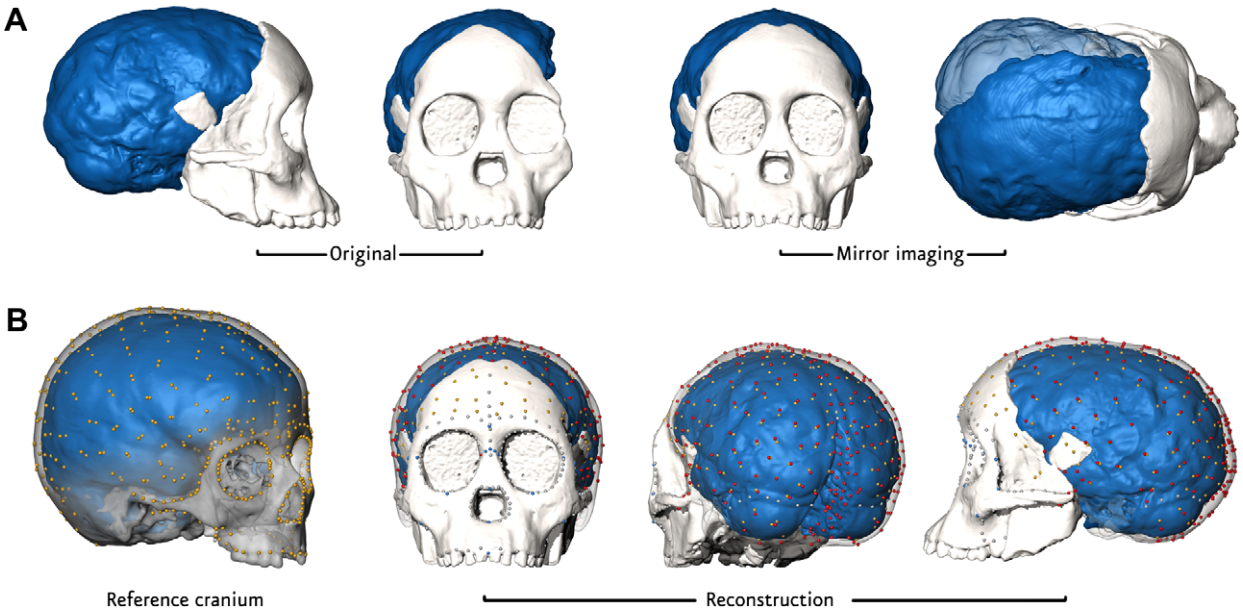

Figure 7. Reconstruction of the Taung cranium. A: The facial and endocranial fragments were translated and rotated so as to exactly align the surfaces of the two parts. Then the less complete left side was replaced by the reflected right side. B: 700 landmarks and semilandmarks were measured on a complete reference cranium (a modern human infant). We used the thin-plate spline interpolation to warp the surface vertices from the reference cranium onto Taung. Anatomical landmarks are drawn in blue, curve semilandmarks in gray, surface semilandmarks in orange, and missing points in red.

Table 1
Anatomical landmarks for the Taung reconstructions.[a]

Prosthion
Nasospinale
Rhinion
Nasion
Glabella
Foramen incisivum
Staphylion
Hormion
Endobregma
Cerebellum L + R
Nasomaxilla L + R
Maxillofrontale Point L + R
Frontomalare orbitale L + R
Frontomalare temporale L + R
Zygorbitale L + R
Zygomaxillare L + R
Jugale L + R

[a] Landmarks in italics were considered semilandmarks after the initial warping step.

The landmarks missing on the Turkana boy after mirror-imaging were estimated by thin-plate spline interpolation, using every complete cranium in our data set as a reference specimen, thereby producing 250 reconstructions of the Turkana boy. Figure 9 shows one of these reconstructions and the set of measurement points.

In a principal components analysis, we treated the several hundred reconstructions as individual specimens to assess the impact of our reference choice on the analysis. Figure 10A shows principal components in Procrustes shape space of the subset of recent and fossil *Homo* crania; Fig. 10B shows the principal components of the complete data set. The 250 reconstructed specimens were projected into the principal components space of the complete crania—otherwise the reconstructions would dominate the principal components analysis. Some variability among the reconstructions is visible, yet the cluster of reconstructed specimens is clearly distinct from the other crania. In Fig. 10A the principal components axes were computed using the *Homo* crania and the average of all reconstructions of KNM-WT 15000; the 250 reconstructions were then projected into this PC space. The differences among the reconstructions are likewise irrelevant in the comparative context of inter-generic shape differences (Fig. 10B). The cloud of reconstructions plots almost on the same spot in the first three principal components of either data set.

In Fig. 10C we show the first three principal components of shape space of just the 250 estimations of KNM-WT 15000 and the respective reconstructed shapes. This is, in effect, a PCA of the reconstruction point cloud in Fig. 10B. Each reconstruction is colored according to the group affiliation of the reference cranium. Naturally, the posterior distribution of the reconstructions reflects the shape differences among the reference crania that served as the prior distribution used to impute the missing data. A cluster of reconstructions based on *Homo* is clearly distinct from a cluster derived from chimpanzee and gorilla reference specimens. These shape differences are visualized in Fig. 10C, where the reconstructed landmarks and semilandmarks are plotted as color-coded wireframes. Clearly, the reconstructed outlines of the orbits and the glabellar region are slightly but consistently different depending on the group of the reference specimen. This means that when using a modern human reference the reconstructed area looks like a modern human; with a chimp as reference, the reconstruction looks like an ape.

Table 2
Semilandmark curves for the Taung reconstructions.[a]

Alveolar curve	Measured following the outer alveolar margin
Nasal aperture	Along the piriform aperture
Orbital rim	Along the inner orbital rim L + R
Torus supraorbitalis	From *frontomalare temporale* to *glabella* L + R
Upper zygomatic	From *auriculare* to *jugale* to *frontomalare temporale* L + R
Lower zygomatic	Inferior margin of zygomatic bone, along the zygomatic root
Palate	Midsagittal from *foramen incisivum* to *staphylion*
Midsagittal exterior	From *glabella* to *opisthion*
Midsagittal interior	From *foramen caecum* to *internal opisthion*
Sinus transversus	Measured along the transversal sinus L + R

[a] Curves in italics are incomplete in Taung.

Discussion

Correcting for deformation

Despite the fairly substantial amount of simulated deformation applied to Sts 5 (Fig. 4), the match between original and reconstruction is almost perfect after symmetrizing the deformed data using RR. The simulations in Fig. 5 demonstrate that this approach makes it possible to correct distorted crania as long as the deformation is fairly uniform across the specimen. An affine deformation cannot be restored perfectly, however, by RR; it always introduces a bias concerning the overall dimensions of the form. This bias is not necessarily small when shears are large. A shear makes a form relatively longer or wider, which cannot be restored by averaging with its reflection, but only by an affine shear again. This shear, however, can only be estimated correctly when the axis of the original shear is known. For the general algebra of concerns like this, see the discussion in Bookstein (1996b). The shears we were exploring here come under the heading of "bilateral mismatch" in Bookstein and Mardia (2003). A specific algorithm for correcting this bias will be the subject of another publication.

In the presence of relatively small affine deformations, RR is an adequate method for restoring symmetry. All the corrected specimens are very close in shape to the original, and all of them fall within the 95% confidence ellipsoid of the same species in the simulations in Fig. 5. These simulations were intended to probe the technique until it breaks down. It is evident in Fig. 5C that the more heavily distorted specimens are farther away in terms of Procrustes distance from the original shape after correction than the ones with only slight deformation.

These simulations suggest that RR can be used to reliably remove the effects of uniform shear for crania that exhibit deformations similar to the simulation in Fig. 4 and Arago XXI in Fig. 6. The bias mentioned above means, in the former case, that the corrected form is a few millimeters wider than the original, which for most applications will be an acceptable tradeoff for being able to include a corrected fossil.

Angielczyk and Sheets (2007) used simulations in 2D to compare different geometric morphometric "retrodeformation" techniques, including RR. They concluded that even though these approaches often yielded visually appealing results, and although the corrected specimens strongly resembled the originals, the reconstructions should nevertheless be used with caution, as reconstructions have an altered covariance structure. That the covariance matrix changes should not come as a surprise because these approaches remove the asymmetry component. For most analyses, this will have little or no impact if only a few corrected fossil specimens are analyzed in the subspace of the first few principal components of complete specimens. Angielczyk and Sheets (2007) artificially deformed and then "retrodeformed" a whole *sample* of specimens and compared its covariance structure to the original covariance structure of the undeformed specimens, whereas we analyzed just a single corrected specimen in the principal components space of the complete crania. While we agree with Angielczyk and Sheets (2007) that one should not use correction techniques on a whole sample, their caveat does not

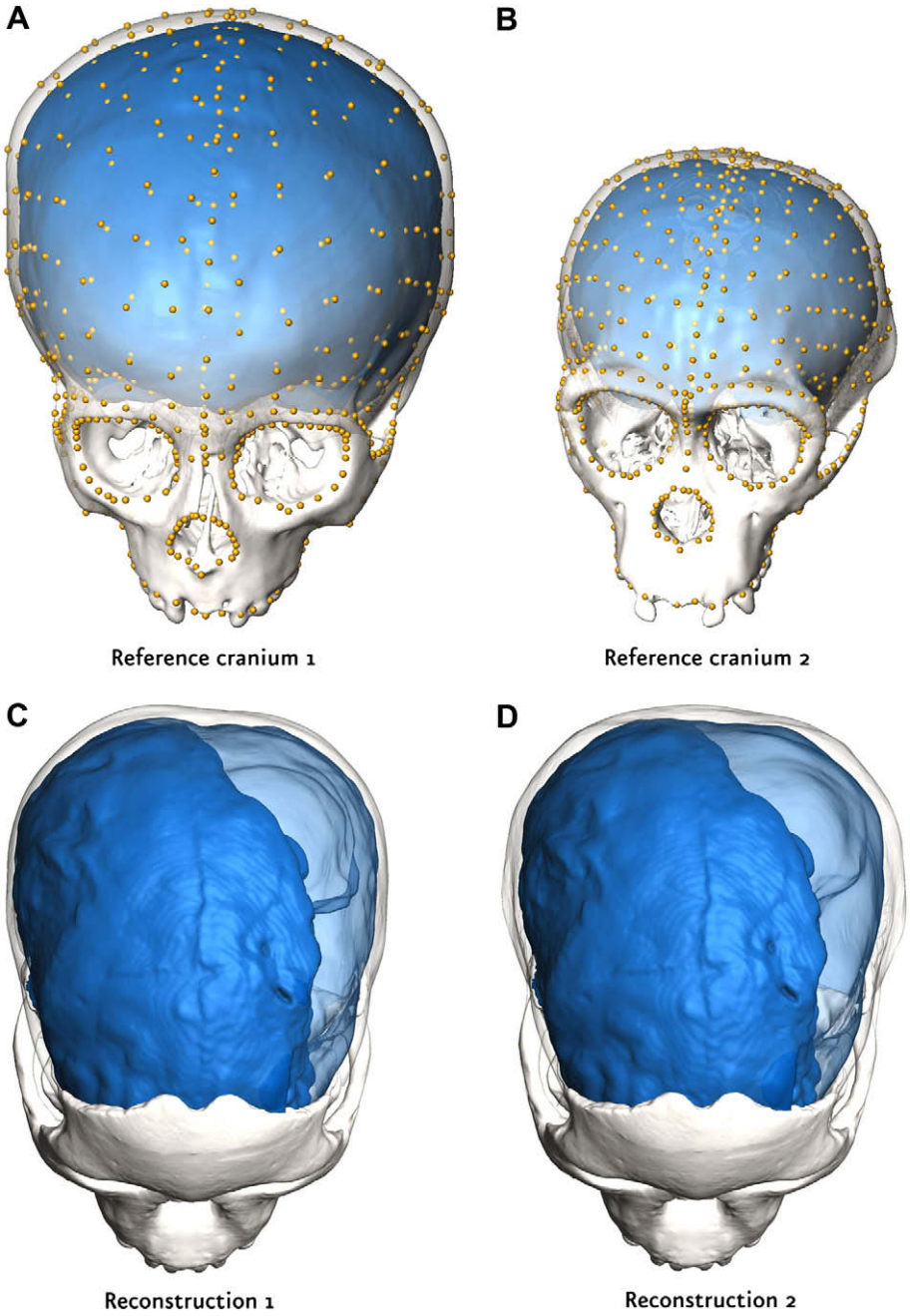

Figure 8. Taung reconstructed twice (see Fig. 7), using two different reference crania (A and B). The left lower panel (C) shows the reconstruction based on a modern human infant; the right lower panel (D) shows the reconstruction using a chimpanzee child as a reference. Despite the morphological differences between the reference crania, the shape differences between the reconstructions are very subtle except in the mastoid region.

apply to analyzing single corrected specimens, provided the sample comprises enough undistorted specimens of the same group.

An alternative approach addressing their concern would symmetrize all specimens of the dataset using reflected relabeling, not just the distorted one, prior to analysis. This would remove the asymmetry component from all specimens and thus make the covariance structures of, for example, a modern reference population and a fossil group that contains corrected specimens, more comparable.

Estimating missing coordinates

Missing data problems do not arise only in fossil material but are common in all morphometric contexts. The sequence of symmetrizing followed by geometric reconstruction advocated here is as useful for estimating missing landmarks in comparative modern series as it is for fossil material, and is applicable, in principle, to any coordinate measurements on bilateral biological structures. Once the missing coordinates are estimated, incomplete specimens can be included in statistical analyses of shape. Hence, the methods presented here are useful for analyzing archaeological and paleoanthropological material. We have demonstrated that if only a few coordinates are missing in the vicinity of preserved anatomy, regression-based statistical reconstruction and TPS-based geometric reconstruction yield very similar results. But even though the shape differences between the different reconstructions are usually very subtle, the statistical properties of the reconstructed sample are

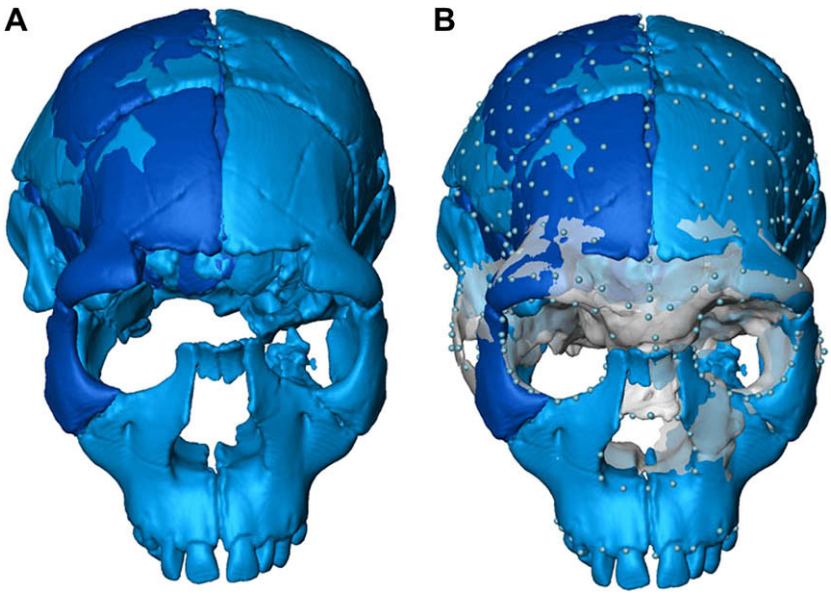

Figure 9. (A) The "Turkana boy" KNM-WT 15000 was first reconstructed by mirror-imaging the undistorted left orbital region (dark blue). (B) 47 landmarks and 300 semilandmarks on the face and the exterior vault were measured on the incomplete fossil and 250 complete crania. Using each of the latter in turn as a reference, the missing points in the supraorbital region were estimated using geometric reconstruction based on the thin-plate spline computed from the non-missing landmarks and semilandmarks. One of the resulting 250 reconstructions is drawn as a semi-transparent gray surface.

different. While geometric reconstruction is based on the smoothness properties of the thin-plate spline, statistical reconstruction is computed using the sample covariance matrix. If the completed forms are to be used in statistical analyses that are computed from the sample covariance matrix, such as regressions, we suggest geometric reconstruction rather than regression-based methods, as in the latter case the linear relationship is "overfitted."

Number of landmarks and semilandmarks

The computational methods discussed here are not applicable to material that is so fragmentary that no landmarks can be identified at all. Although algebraically only four landmarks must be present on a fragment for a three-dimensional thin-plate spline to be computable, a specimen should have *far* more than that. Also, estimation via thin-plate spline interpolation should be used only when missing data are estimated in close proximity to actually observed landmarks and semilandmarks. The thin-plate spline is designed to be multi-local. Therefore, extreme procedures, such as estimating landmarks on the occipital bone when only facial landmarks are present, cannot succeed. Obviously, in such an extreme case, preserved landmarks fail to bend the spline in the vicinity of the missing landmarks in any meaningful way.

All of our examples have therefore used semilandmarks on curves and surfaces in addition to traditional anatomical landmarks. The statistical and geometric algorithms described here cannot magically create information where there is none. To reconstruct missing parts of a skull reliably, the unobservable coordinates have to be predictable, either via the covariance matrix of a reference sample or via the smoothness properties of the thin-plate spline. We define reconstruction accuracy as the mean-squared difference between the original and the reconstruction in appropriate units (e.g., squared Procrustes distance). Geometric reconstruction via thin-plate spline can reliably predict the missing data only if there exists measured coordinates close to the defect. Statistical reconstruction via regression, on the other hand, relies on the covariation among the observable coordinates. As closely spaced points tend to have a higher covariance than more distant landmarks, the use of semilandmarks greatly improves the accuracy of the regression-based algorithms as well.

Using simulations (Gunz et al., 2004, submitted for publication; Gunz, 2005) we have shown that the accuracy of estimation depends on the number of missing coordinates, the morphology of the missing part, its variability in the reference population, and the point spacing and distribution of the coordinates on the non-missing parts. After deliberately deleting landmarks and semilandmarks digitized on complete modern human crania, we compared the reconstructions with their originals. In these simulations we found that geometric reconstruction via TPS performed better, on average, than regression-based methods and mean substitution. Estimating missing data via regression requires a sufficiently large reference sample; otherwise the error in estimating the regression parameters outweighs its advantages. As covariance matrices require a larger sample to be reliably estimated than do means, geometric reconstruction via TPS can also be applied to a smaller sample.

The missing data methods described here exploit the fact that the information captured by shape coordinates typically tends to be highly redundant, especially when the measurement points are closely spaced. Reconstruction can work only if the morphology is covered sufficiently by measurements. When the number of coordinates is too low, no reliable covariance structure exists among the landmarks (hence, no regression is possible) and a TPS deformation may not be meaningful. Therefore, these methods are not very well suited for datasets that use anatomical point-landmarks only. Whether they were measured using a coordinate digitizer such as a Microscribe (Immersion, Inc.) or on surface or CT scans, it is the density of measurement points that matters most for the accuracy of estimation. For most regions of the skull, the dense point spacing required for estimating data that are missing on both sides or in the midplane is only achievable by using semilandmarks. Reflection of coordinates from one side to the other, in contrast, can be used when only anatomical landmarks are available. Similarly, correcting for fairly uniform shears via RR works well with only landmark data.

As we have already mentioned, it is impossible to give a general recommendation about the number of missing landmarks that can

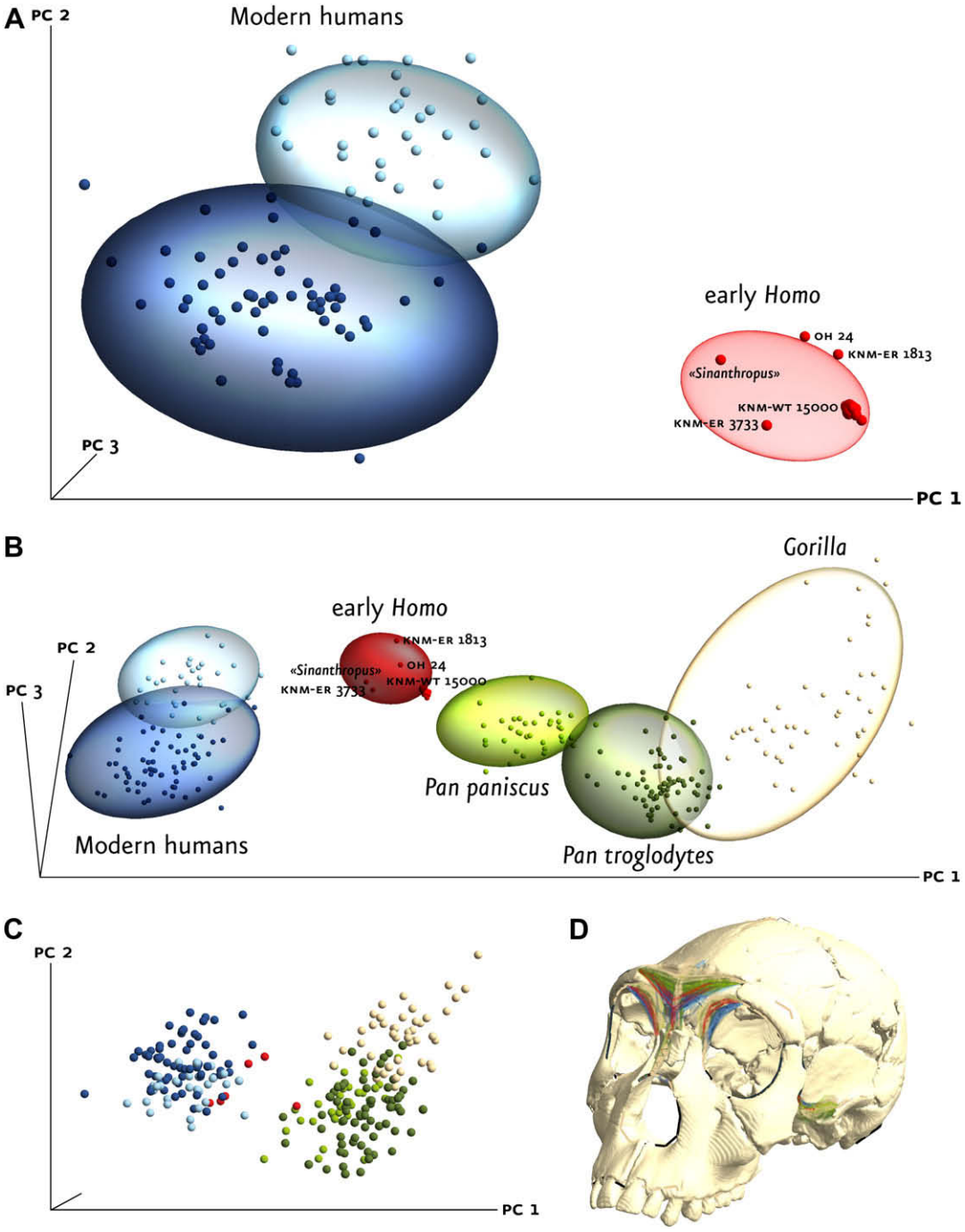

Figure 10. (A) Principal components (PC) 1–3 in shape space of the subset of recent and fossil *Homo* crania. (B) PCA of the complete data set. The cloud of reconstructions plots almost on the same spot in the first three principal components of either data set. The 250 reconstructed specimens were projected into the PC space of the complete crania (otherwise the reconstructions would dominate the principal components analysis). (C) PCA of just the 250 estimations of KNM-WT 15000 and the respective reconstructed shapes. Each reconstruction is colored according to the group affiliation of the reference cranium. A cluster of reconstructions based on *Homo* is clearly distinct from a cluster derived from chimpanzee and gorilla reference specimens. The reconstructed landmarks and semilandmarks are plotted as color-coded wireframes (D). Clearly, the reconstructed outlines of the orbits and the glabellar region are slightly but consistently different depending on the group of the reference specimen.

be estimated or the percentage of the cranium that can be missing, as such numbers would vary greatly between different regions of the organism. Figure 7, the reconstruction of the Taung cranium, and Fig. 9, the reconstruction of the glabellar area of KNM-WT 15000, offer good examples of how close actually measured points should be in the vicinity of the missing coordinates in order for predictions of the missing data to be reliable. The quality of the estimate also depends on the nature of the biological variation. The outer neurocranial shell in Taung is tightly integrated with the inner shell, and for this reason, can be estimated with high precision given the available data from the endocast. The morphology of the glabellar area is variable across species, and also within species, partly due to its late development (Mitteroecker and Bookstein, 2008). This renders it more difficult to estimate from the available morphology than the much higher number of missing coordinates in the Taung cranium.

The difference between Fig. 2A and B illustrates why our results and recommendations are different from the ones of Neeser et al. (in press). Those authors simulated random-knockout of points using a dataset of 29 landmarks and concluded that regression-based methods outperformed TPS-based estimation when using only a few anatomical landmarks. The 29 anatomical landmarks used in Neeser et al. (in press) and Fig. 2 do not capture the morphology of the reference and target specimen in sufficient detail to make a reliable prediction of missing coordinates. In this case, mean substitution gives the "best" least squares estimate (that is, after all, the definition of the mean). But mean substitution is not a reconstruction method. If mean substitution performs best in a simulation, this indicates that missing data estimation should not be used at all.

Note also that Neeser et al. (in press) dealt not with shape coordinates, but with measured distances. The number of distances is, of course, vastly greater than the number of degrees of freedom of the data—for the 29 landmarks of the Neeser et al. (in press) example, there are 435 distances, compared to only 55 degrees of freedom for the full size space. The multiplicity of regression estimates for locations of what is after all the same point must be adjusted by methods of multidimensional scaling that have nothing to do with the statistical models for landmark variation, Procrustes distance, or any other component of the morphometric toolkit. In short, there is no reason to expect any agreement between the computations of Neeser et al. (in press) and the principled methods of this paper and its companions (Gunz et al., 2004, submitted for publication). There is no comparably powerful methodology for the reconstruction of missing landmarks when shape or form is represented by a matrix of Euclidean distances (EDMA). Our TPS algorithm is applied to the landmark coordinates directly instead of to interlandmark distances. While two landmark configurations need not be superimposed in order to compute a TPS interpolation, missing data estimation based on a group mean form as reference necessarily involves a Procrustes analysis. Also, the regression approach cannot be applied to Euclidean distance matrices equally well, as the relationship among such distances is neither linear nor exponential. Averaging a form with its reflection involves a complex rearrangement of the distance matrices and the ensuing average distances do not correspond to a valid landmark configuration, which can only be approximated.

Which semilandmarks are missing?

One of the crucial steps in the reconstruction protocol is to establish exactly which coordinates have to be estimated. This is easy for anatomical landmarks because their absence can be recorded during digitizing. If the measurement protocol includes semilandmarks, then the only trivial case is a wholly absent curve or surface. If parts of a measured structure are missing, then defining the exact break point of the defect typically includes at least some manual intervention, as illustrated in our examples.

Exploiting available morphology

In the geometric reconstruction algorithm, one point-set creates a TPS mapping function, while another set of points, located on the reference form, is mapped onto the incomplete target form. These point-sets need not be the same. In the Taung reconstruction a data-set of 700 landmarks and semilandmarks including endocranial information was used for reconstruction, whereas a different set of 347 points was later used in the analysis (Gunz, 2005). A much higher resolution of coordinates was therefore used for the reconstruction than for the subsequent analyses, so as to exploit as much of the available geometry as possible.

Choosing a reference specimen

The choice reference specimen for TPS-based estimation has a large impact on the resulting reconstruction. Especially for fossil taxa that are represented by only a few specimens, the choice of reference specimens is crucial. If data and signals are so vague or ambiguous that inferences depend more on assumptions than on the actual data, then throughout the analysis we need to keep track of the uncertainties due to the missing data values and the sensitivity to the prior assumptions. One can think of the distribution of reference crania as a Bayesian prior distribution, and of the reconstructed specimens as the corresponding posterior distribution. This posterior distribution reflects the uncertainty due to the missing data values and the sensitivity to the prior assumptions.

In our reconstruction of the Taung cranium, the differences between the human and the chimpanzee model underlying the reconstruction were visible only in the mastoid region. The mastoid region in the chimpanzee-based reconstruction was more inflated than in the human-based estimation because (1) there were only few preserved landmarks and semilandmarks in the vicinity, and (2) the reference crania differ substantially in their mastoid morphology (Fig. 8). Similarly, the posterior distribution of the reconstructions of KNM-WT 15000 (Fig. 10C) showed two distinct clusters. The posterior distribution of reconstructions like this *must* reflect the prior distribution of reference specimens.

Which reference cranium should one pick then, under which circumstances? The answer depends on the context in which the reconstruction will be used and in the following we will sketch four scenarios for which we feel comfortable giving recommendations.

In a large series of extant comparative material, usually only a few coordinates are missing on some specimens and there exists a sample of complete reference specimens to which the incomplete forms belong. In such cases, one can estimate the missing coordinates by exploiting symmetry information first and then via geometric reconstruction, using the group mean of the complete shapes as the reference shape. The data matrix of completed forms can then often be analyzed as if there had been no missing coordinates in the first place, inasmuch as the impact of the estimation algorithms is negligible.

Figure 10 represents a typical application of geometric morphometrics in paleoanthropology: the shape variation of a fossil group (early *Homo*) is compared in an inter-specific (Fig. 10A) and an inter-generic framework (Fig. 10B) using principal components analysis. We have shown that the choice of reference specimen is almost inconsequential for the dimensions of shape space that were interpreted. Even though a pattern within these different reconstructions is apparent (Fig. 10C), the shape differences among the reconstructions are too small to alter any inferences drawn from the PCA analysis—even when using reference crania that are obviously inappropriate.

If we wanted to create a stereolithographic model of the Turkana-Boy reconstruction for a museum, it would make sense to use another *Homo erectus s.l.* cranium as a reference, or a mean *Homo erectus s.l.* form computed as a Procrustes average. When the surface of the incomplete specimen is merged with such a reconstructed surface, the final form thus completed can also be modeled using rapid prototyping technology.

If the scientific aim is to explore taxonomic affinities of an incomplete specimen, one must be careful to avoid circular reasoning, such as would be the case if one reconstructed missing data in Arago XXI using a Neanderthal reference specimen and then claim its phenetic affinities with Neanderthals. One would also want to create reconstructions using a modern human reference model to support a claim along the lines of "even if reconstructed using modern human reference crania, it clusters with Neanderthals."

(This is just hypothetical; we are actually claiming no such finding.) In a similar vein, we reconstructed the exterior shell of the Taung child using a chimpanzee and a human model. Because the shape differences between these two "extremes" were not sufficient to alter the conclusions drawn from the analysis, we considered the reconstruction to be reliable in those aspects that were analyzed in Gunz (2005). We have shown, however, that using a chimpanzee model to reconstruct Taung creates an ape-like morphology in the mastoid region; this was irrelevant in the aforementioned analysis of Gunz (2005), but might matter for a different scientific question.

Conclusions

As soon as parts of a specimen are missing or deformed, *any* reconstruction—using either computer algorithms or sculpting with plaster—must make assumptions about functional constraints, integration and symmetry, and sometimes also about sex, taxonomic affinity, and taphonomy. The methods presented here require explicit recognition of the assumptions behind every reconstruction, so that reconstruction validity can subjected to evaluation and discussion. Whether or not similar conclusions follow from a variety of realistic alternative assumptions can thus be checked during analysis. Because different assumptions and algorithms lead to different estimations, there exists no "all-purpose" reconstruction. Instead, one creates multiple reconstructions—a posterior distribution in a Bayesian sense, for which the reference specimens serve as the prior distribution. For crucial fossil specimens we therefore suggest creating multiple plausible reconstructions. For example, one could use all complete crania as reference models (e.g., Fig. 10).

A great advantage of geometric morphometrics in general is that all results can be visualized, so it is very easy to check the results of the reconstruction algorithms visually. Applying geometric reconstruction to incomplete fossil crania for which only a few complete reference specimens exist requires caution at every step of the reconstruction. We recommend visually checking all stages of the reconstruction protocol by warping the surface of the reference specimen to the reconstructed configuration (Figs. 3 and 7–9).

Attempts to address reconstruction uncertainty should be an essential part of any analysis that includes completed forms. Here we have shown one way to visualize this uncertainty—creating a distribution of reconstructions (Fig. 10). This distribution does not primarily reflect sampling error as in most standard statistical contexts, but mainly the uncertainty due to the arbitrary choice of references and biological/paleontological assumptions. The same logic of multiple virtual reconstructions, albeit manual ones, was applied in Zollikofer et al.'s (2005) virtual reconstruction of the distorted *Sahelanthropus tschadensis* cranium TM 266. These authors used two different reconstruction protocols, manually aligning and un-distorting bone fragments while constraining the overall morphology to be symmetric. The uncertainty of estimated data matters not only for morphometric analysis but also for other models such as finite element computations (e.g., Strait et al., 2009). While there typically will be shape differences among equally plausible reconstructions, these different estimates might still support a single conclusion. But they need not do so, and all assumptions must be strenuously challenged if one or more reconstructions, or a statistical analysis based on them, are to be treated as arguments for a scientific claim.

Acknowledgements

We thank the curators of fossil or extant specimens for access to specimens in their care. The extant hominoid sample comprises crania housed in the following institutions: the Department for Anthropology, University of Vienna, Austria, and the Institute for Anthropology and Human Genetics, Frankfurt/Main, Germany. We thank Francis Thackeray and Stephany Potze from the Transvaal museum for their hospitality and help, and likewise Mike Raath, Lee Berger, Phillip Tobias, from Johannesburg, as well as Emma Mbua in Nairobi. We also want to thank K. Kuykendall, J.H. Labuscagne, and Fred Spoor for their support regarding the CT scans of some of the fossil specimens. We are indebted to Markus Bernhard and Caroline Fenes for sharing their dataset of extant primates. We want to thank Katerina Harvati, Jean-Jacques Hublin, Katrin Schäfer, Horst Seidler, and Andrea Stadlmayr for their support and advice.

Our research was supported by the FWF grant P14738, Austrian Council for Science and Technology grant GZ.200.093/I-VI/I/2004, by the EU FP6 Marie Curie Actions grant MRTN-CT-2005-019564 "EVAN" and the Max Planck Society.

References

Adams, D., Rohlf, F.J., Slice, D.E., 2004. Geometric morphometrics: Ten years of progress following the 'revolution'. Italian J. Zool. 71, 5–16.
Angielczyk, K.D., Sheets, D., 2007. Investigation of simulated tectonic deformation in fossils using geometric morphometrics. Paleobiology 33, 125–148.
Bookstein, F.L., 1991. Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge University Press, Cambridge.
Bookstein, F.L., 1996a. Combining the tools of geometric morphometrics. In: Marcus, L.F., Corti, M., Loy, A., Naylor, G., Slice, D.E. (Eds.), Advances in Morphometrics. Plenum Press, New York, pp. 131–151.
Bookstein, F.L., 1996b. Standard formula for the uniform shape component in landmark data. In: Marcus, L.F., Corti, M., Loy, A., Naylor, G., Slice, D.E. (Eds.), Advances in Morphometrics. Plenum Press, New York, pp. 153–168.
Bookstein, F.L., 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Med. Image Anal. 1 (3), 225–243.
Bookstein, F.L., 2005. After landmarks. In: Slice, D.E. (Ed.), Modern Morphometrics in Physical Anthropology. Kluwer Academic/Plenum Publishers, New York, pp. 49–71.
Bookstein, F.L., Green, W.K., 2002. Edgewarp 3D Available from: ftp://brainmap.stat.washington.edu/pub/fred/ewsh3.19.man/.
Bookstein, F.L., Mardia, K.V., 2003. The five components of directional asymmetry. In: Aykroyd, R.G., Mardia, K.V., Langdon, M.J. (Eds.), Stochastic Geometry, Biological Structure, and Images. Dept. of Statistics, University of Leeds, pp. 35–40.
Bookstein, F.L., Gunz, P., Mitteroecker, P., Prossinger, H., Schaefer, K., Seidler, H., 2003. Cranial integration in *Homo*: singular warps analysis of the midsagittal plane in ontogeny and evolution. J. Hum. Evol. 44 (2), 167–187.
Broom, R., 1947. Discovery of a new skull of the Sterkfontein Ape man, *Plesianthropus*. Nature 159, 672.
Conroy, G.C., Vannier, M.W., 1984. Noninvasive three-dimensional computer imaging of matrix-filled fossil skulls by high-resolution computed tomography. Science 226 (4673), 456–458.
Conroy, G.C., Vannier, M.W., 1987. Dental development of the Taung skull from computerized tomography. Nature 329 (6140), 625–627.
Dart, R.A., 1925. *Australopithecus africanus*: the man-ape of South Africa. Nature 115, 195–199.
Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood estimation from incomplete data via the EM algorithm. J R Stat Soc Series B 39, 1–38.
Gunz, P., 2005. Statistical and geometric reconstruction of hominid crania: reconstructing australopithecine ontogeny. Ph.D. Dissertation, University of Vienna.
Gunz, P., Harvati, K., 2007. The Neanderthal "chignon": Variation, integration, and homology. J. Hum. Evol. 52 (3), 262–274.
Gunz, P., Mitteroecker, P., Bookstein, F.L., 2005. Semilandmarks in three dimensions. In: Slice, D.E. (Ed.), Modern Morphometrics in Physical Anthropology. Kluwer Academic/Plenum Publishers, New York, pp. 73–98.
Gunz, P., Mitteroecker, P., Bookstein, F.L., Weber, G.W., 2004. Computer Aided Reconstruction of Incomplete Human Crania using Statistical and Geometrical Estimation Methods. Enter the Past: Computer Applications and Quantitative Methods in Archaeology. BAR International Series, Oxford, pp. 96–98.
Gunz, P., Mitteroecker, P., Weber, G.W., Bookstein, F.L. Estimating missing data using geometric morphometrics. Syst. Biol., submitted for publication.
Jungers, W.L., Minns, R.J., 1979. Computed tomography and biomechanical analysis of fossil long bones. Am. J. Phys. Anthropol. 50 (2), 285–290.
Kalvin, A.D., Dean, D., Hublin, J.J., 1995. Reconstruction of human fossils. IEEE Comput. Graph Appl. 15, 12–15.
de Lumley, H., de Lumley, M.-A., 1971. Découverte de restes humaines anténéandertaliens datés au début de Riss à la Caune d'Arago (Tautavel, Pyrénées-Orientales). C.R. Acad. Sci. Paris 272, 1729–1742.
Mardia, K.V., Bookstein, F.L., Moreton, I.J., 2000. Statistical assessment of bilateral symmetry of shapes. Biometrika 87, 285–300.

McLachlan, G.J., Krishnan, T., 1997. The EM Algorithm and Extensions. Wiley, New York.
Mitteroecker, P., Bookstein, F.L., 2007. The conceptual and statistical relationship between modularity and morphological integration. Syst. Biol. 56 (5), 818–836.
Mitteroecker, P., Bookstein, F.L., 2008. The evolutionary role of modularity and integration in the hominoid cranium. Evolution 62 (4), 943–958.
Mitteroecker, P., Gunz, P., 2009. Advances in Geometric Morphometrics. Evol. Biol. 36 (2), 235–247.
Motani, R., 1997. New technique for retrodeforming tectonically deformed fossils, with an example for ichthyosaurian specimens. Lethaia 30, 221–228.
Neeser, R., Ackermann, R.R, Gain, J. Comparing the accuracy and precision of three techniques used for estimating missing landmarks when reconstructing fossil hominin crania. Am. J. Phys. Anthropol., in press. doi: 10.1002/ajpa.21023.
Petersen, H.C., Ahlström, T., Vach, W., 2006. Geometric morphometry, missing values, and multiple imputations. Am. J. Phys. Anthropol. S42, 91.
Ponce de León, M.S., 2002. Computerized paleoanthropology and Neanderthals: the case of Le Moustier 1. Evol. Anthropol. 11 (Suppl. 1), 68–72.
Ponce de León, M.S., Zollikofer, C.P.E., 1999. New evidence from Le Moustier 1: computer-assisted reconstruction and morphometry of the skull. Anat. Rec. 254 (4), 474–489.
Rak, Y., 1983. The Australopithecine Face. Academic Press, New York.
Rohlf, F.J., 1993. Contributions to morphometrics. In: Marcus, L.F., Bello, E., García-Valdecasas, A. (Eds.), Museo Nacional de Ciencies Naturales, Madrid, pp. 131–159.
Schafer, J.L., 1997. Analysis of Incomplete Multivariate Data. Capman and Hall, London.
Schafer, J.L., Graham, J.W., 2002. Missing data: our view of the state of the art. Psychol. Methods 7 (2), 147–177.
Slice, D.E., 2007. Geometric morphometrics. Ann. Rev. Anthropol. 36, 261–281.
Strait, D.S., Weber, G.W., Neubauer, S., Chalk, J., Richmond, B.G., Lucas, P.W., Spencer, M.A., Schrein, C., Dechow, P., Ross, C.F., Grosse, I., Wright, B.W., Constantino, P., Wood, B., Lawn, B., Hylander, W.L., Wang, Q., Byron, C., Slice, D.E., Smith, A.L., 2009. The feeding biomechanics and dietary ecology of *Australopithecus africanus*. Proc. Natl. Acad. Sci. U.S.A 106 (7), 2124–2129.
Tate, J.R., Cann, C.E., 1982. High-resolution computed tomography for the comparative study of fossil and extant bone. Am. J. Phys. Anthropol. 58 (1), 67–73.
Tattersall, I., Sawyer, G.J., 1996. The skull of "Sinanthropus" from Zhoukoudian, China: a new reconstruction. J. Hum. Evol. 31 (4), 311–314.
Vannier, M.W., Conroy, G.C., Marsh, J.L., Knapp, R.H., 1985. Three-dimensional cranial surface reconstructions using high- resolution computed tomography. Am. J. Phys. Anthropol. 67 (4), 299–311.
Walker, A., Leakey, R., 1993. The Nariokotome *Homo erectus*, Skeleton. Harvard University Press, Cambridge.
Weidenreich, F., 1937. Reconstruction of an entire skull of an adult female individual of Sinanthropus pekinensis. Nature 140, 1010–1011.
Wind, J., 1984. Computerized X-ray tomography of fossil hominid skulls. Am. J. Phys. Anthropol. 63 (3), 265–282.
Zollikofer, C.P.E., Ponce de León, M.S., Lieberman, D.E., Guy, F., Pilbeam, D., Likius, A., Mackaye, H.T., Vignaud, P., Brunet, M., 2005. Virtual cranial reconstruction of *Sahelanthropus tchadensis*. Nature 434 (7034), 755–759.
Zollikofer, C.P.E., Ponce De León, M.S., Martin, R.D., 1998. Computer-assisted paleoanthropology. Evol. Anthropol. 6 (2), 41–54.
Zollikofer, C.P.E., Ponce de Leon, M.S., Martin, R.D., Stucki, P., 1995. Neanderthal computer skulls. Nature 375 (6529), 283–285.
Zollikofer, C.P.E., Ponce de León, M.S., Vandermeersch, B., Léveque, F., 2002. Evidence for interpersonal violence in the St. Césaire Neanderthal. Proc. Natl. Acad. Sci. U.S.A 99 (9), 6444–6448.