

COMPUTER AIDED RECONSTRUCTION OF HUMAN CRANIA

ABSTRACT

P. GUNZ

VIENNA INST. FOR ANTHROPOLOGY, VIENNA, AUSTRIA

P. MITTEROECKER

VIENNA INST. FOR ANTHROPOLOGY, VIENNA, AUSTRIA

FL. BOOKSTEIN

VIENNA INST. FOR ANTHROPOLOGY, VIENNA, AUSTRIA
MICHIGAN CENTER OF BIOLOGICAL INFORMATION,
UNIV. OF MICHIGAN, USA

GW. WEBER

VIENNA INST. FOR ANTHROPOLOGY, VIENNA, AUSTRIA

We present statistical and geometrical techniques to reconstruct incomplete human crania using techniques that formalize the biologist's prior understanding of the considerations that govern form: continuity, symmetry and integration. The modern morphometrics of landmarks and curves makes it possible to blend statistical and biological reasoning in this domain. Factors such as size allometry or sex and also directional asymmetry whether zero or nonzero can be explicitly incorporated into the data estimation by way of the corresponding covariance structures. For tasks of estimation based on very small samples we show a variant based on the continuity assumption of the thin-plate spline. When complete specimens are adequate in number our estimation can be regression-driven instead. All the missing points can be estimated at once by maximizing the likelihood of the resulting configuration in a reduced-rank model of a multivariate Gaussian distribution. Whatever integration the form possesses is automatically exploited in the course of these regressions. We demonstrate the accuracy of these approaches using a dataset of 388 anatomical landmarks and semi-landmarks on 52 complete *H. sapiens* crania. After deliberately deleting regions of landmarks we estimate the missing data and compare the estimated specimens to the originals. Our results indicate that the accuracy of estimation is sufficiently close to the precision of measurement.

Life after death is not kind to paleo-anthropological specimens. Forms are distorted, parts break off, and the textural cues we use to locate landmark points may be effaced or covered by matrix. Distortion shifts landmarks away from their original context; we will not discuss that form of data damage here. The effect of the other two types of processes, breakage and effacement or encrustation, is to render the landmark locations actually unobservable in the single specimen; such points are coded as missing. The topic of the present paper is the subsequent treatment of points coded as unobservable in this way.

Across the general run of modern computational statistics there is one core algorithm for handling missing data, the EM ALGORITHM (Dempster et al. 1977). E stands for EXPECTATION, and M for MAXIMISATION. In any EM analysis, an incomplete data set is treated as having arisen from a complete data set by random knockout of observed values, and the data set is completed as a whole (i.e., not specimen by specimen but all at once) in order to maximize some overall probabilistic likelihood, such as a multivariate Gaussian one. The algorithm is iterative: a covariance structure is approximated, then missing data is estimated by regressions datum by datum, then the covariance structure recomputed, the regressions redone, and so on until convergence. The procedure is robust and reliable in many settings within the natural and the social sciences (Allison 2001, Little and Rubin 2002).

However, anthropological questions involving missing data do not necessarily suit any of these typical settings. Data might not be missing at random: Geologically older specimens might be missing more of their landmarks, infant or small individuals have fragile bones that break easily - and if

not the discrete landmark points then certainly the semi-landmarks (Bookstein 1997, Mitteroecker et al., in press) that are missing tend to cluster on forms. We may not be intending to optimize any sort of a likelihood for the sample as a whole, but only some descriptive functional for the reconstruction of one single form; and the purpose of the reconstruction might be to understand gross aspects of size and shape, or instead details of local modeling.

The difference between a concern for spatial position per se, versus a concern for aspects of shape, is closely related to the distinction between two quantities that are familiar separately in the literature of geometric morphometrics: Procrustes distance (sum of squared separations of shape coordinates regardless of spatial position) and bending energy (weighted sum of partial warp scores that take spatial contiguity into account in a very fundamental way) (Bookstein 1991, 1997, Rohlf and Slice 1990).

We will show that either of these concerns can be set at the core of an algorithm for estimating missing data that yields reasonable results in realistic simulations. One approach is using the thin plate spline for GEOMETRIC RECONSTRUCTION while the other uses multiple multivariate regression for STATISTICAL RECONSTRUCTION. The choice between the methods is not, at root, a matter of algebra, but of science: the methods are pertinent to two different paleo-anthropological contexts. We will return to this distinction in the final discussion, after notating and demonstrating the two methods.

At this point we want to completely dismiss a method that is often found in the literature: the method of MEAN SUBSTI-

TUTION borrowed from the social sciences. In a mean substitution, missing entries are substituted by the mean of the non-missing entries over the full sample data set. When data are Cartesian coordinates or shape coordinates this procedure makes no sense either as statistics or as science - because of the method's prevalence, the test results of mean substitution are shown nevertheless for comparison.

THE METHODS

Imagine a data set of many landmarks or semi-landmarks over many specimens. Consider, first, the very simplest case: exactly one landmark is missing for only one specimen. There is a good deal of information available to help us to reasonably estimate its location. In one approach, the STATISTICAL RECONSTRUCTION, we think of this point as correlated with all the other landmark locations of the landmark set. Using the other - complete - specimens, we work out the prediction function that predicts it with the minimum sum-of-squares given the other data.¹

But we might just as well predict the missing parts based on geometric properties of the single specimen, such as continuity information of curvature - GEOMETRIC RECONSTRUCTION. For this purpose we use the thin plate spline interpolation function: We predict the missing data mapping the average of the complete cases to the specimen with missing landmarks - using the thin plate spline interpolation based on the subset of observable landmarks.

These two alternatives yield different locations, in general, but for reasonable numbers of reasonably distributed landmarks the discrepancy is virtually undetectable.

If more than one form is missing landmarks, then whichever figure of merit we choose, we need an iterative approach, because in practical applications not all specimens are missing the same landmarks. First we estimate the missing points, but then using the regression method we have to re-compute the covariance matrix that gives us the prediction formulas (because now it is using all the forms, as they have all been tentatively completed), and in the thin-plate method we have to re-compute the average form that gives us our spline, because now it averages over the entire sample combining both the originally complete forms and the tentatively completed. In either version, the iteration converges rather quickly for reasonable data schemes.

At the conclusion of either of these algorithms, we have estimated all of the missing data anywhere in the data set in order to optimize the quantity that is encoded in the estimation step.

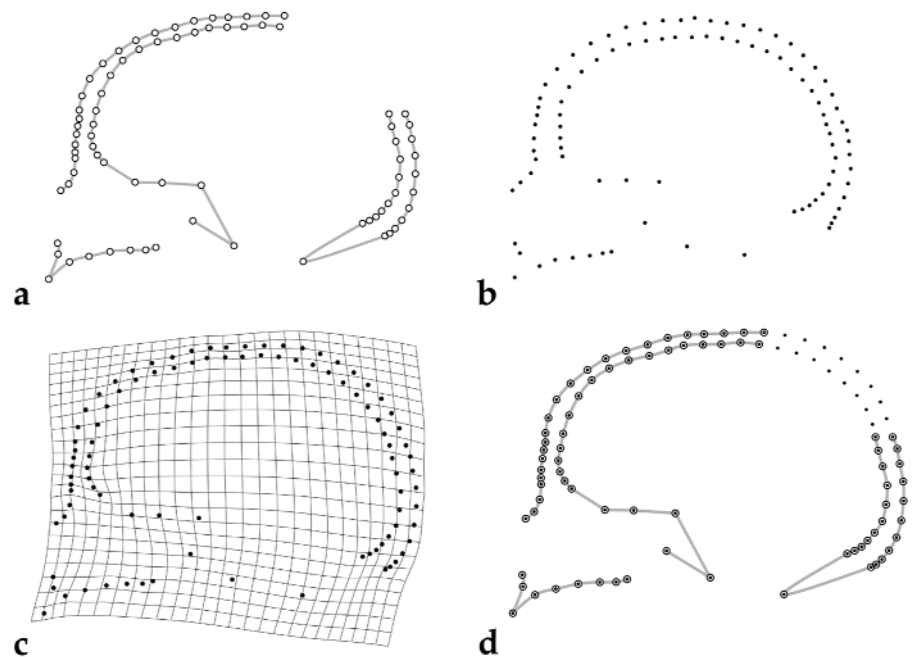


Figure 1 GEOMETRIC RECONSTRUCTION. (a) The target specimen is missing landmarks in the neurocranium. (b) Complete reference specimen. (c) The thin plate spline deformation grid between the specimens in (a) and (b) is computed from the subset of all landmarks available in (a). (d) Missing data of the target specimen is substituted according to the deformation grid shown in (c), which maps the landmark locations of the reference specimen to the target specimen

For the first option, that quantity is the net regression residual sum-of-squares for prediction of landmarks by the others; for the second option, it is the net bending energy of the completed sample around its grand mean.

SIMULATIONS

We demonstrate the accuracy of these approaches using a dataset of 388 anatomical landmarks and semi-landmarks on 52 complete *H. sapiens* crania. After deliberately deleting regions of landmarks we estimate the missing data and compare the estimated specimens to the originals. Figure 1 shows the 4 different KNOCKOUT-SETS that were used to compare the accuracy of the two estimation methods.

First the landmarks in the shaded regions were deleted on one of the 52 specimens, then this specimen was reconstructed using the two methods introduced above. This was done for each individual. As in each calculation-cycle only one form is missing landmarks, there is no need for iteration here. Figure 2 summarizes the results as the mean of the summed squared residuals per landmark for each knockout-set and estimation method. A semi-landmark carries only shape-information perpendicular to the curvature, so when a semi-landmark was deleted, only the residual normal to the ridge or surface was used.

Each bar's length represents the total error of estimation. Mean substitution always performs worst. With the exception of the first knock-out set, the regression method is always better than the thin plate spline warping.

These differences result from intrinsic properties of the compared methods: The spline performs best only in the case of

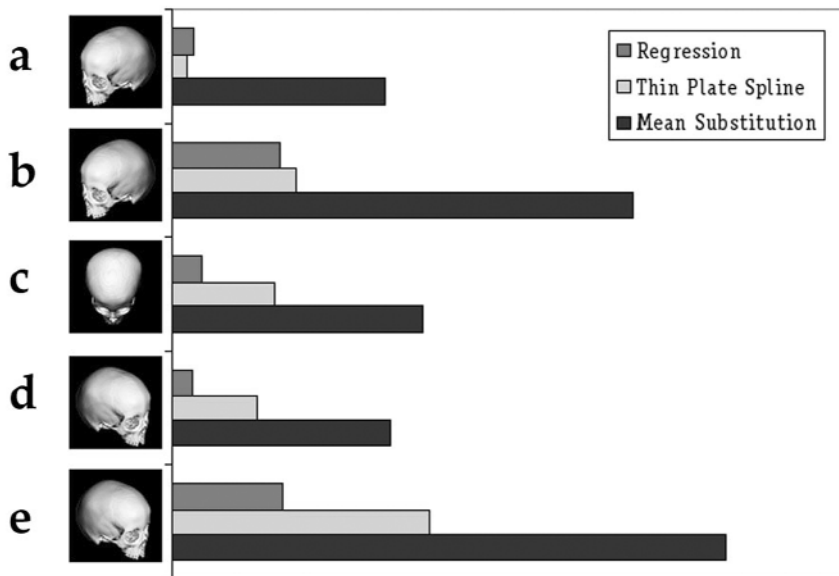


Figure 2 plate spline warping and regression and mean substitution. Each bar's length represents the total error of estimation. Mean substitution always performs worst. With the exception of the first knock-out set (a), the regression method is always better than the thin plate spline warping

Figure 2a, where only a small part of the neuro-cranial surface is missing. The thin plate spline computes the deformation that is least bent, which turns out to be the best method for the smooth curvature of the neuro-cranial surface. Particularly elucidating is the contrast between Figure 2b, where regression and thin plate spline perform almost equally well and 2d and 2e, where the regression is more than twice as precise as the spline. In the former case, the whole face has to be estimated, in the latter, information is missing on only one half of the cranium. While it is equally hard for both methods to reconstruct the face when only neuro-cranial information is available, the different results for the knockout-set of 2d and 2e demonstrate that only the regression has access to symmetry information. But not only symmetry - the regression exploits all information about biological factors present in the reference population like allometry or morphological integration (Bookstein et al. 2003).

PRINCIPAL COMPONENTS

We performed principal components analysis (PCA) of the Procrustes coordinates (Rohlf 1993) of the original and the estimated specimens. Figure 3 shows the first two PC's of the

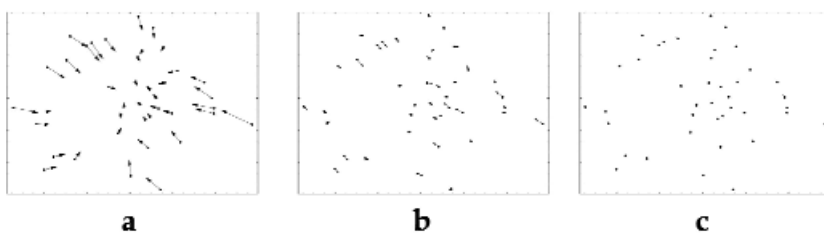


Figure 3 First two principal components of shape of the 52 complete original and the reconstructed specimens for the knockout-set of Figure 2e. The vectors point from the original towards the reconstruction. (a) Mean substitution. Note that the reconstructions are biased towards the mean. (b) Thin plate spline. No obvious bias. (c) Multiple multivariate regression. The error is almost too small to be drawn.

knockout-set of Figure 2e for each method. Each complete and reconstructed specimen is connected with a small vector, where the arrowhead points towards the reconstruction. The arrows of the thin plate spline reconstructions exhibit no obvious patterning and the estimation errors of the regression are almost too small to be seen in the plot of the first two PC's.

DISCUSSION

As we have already argued, MEAN SUBSTITUTIONS should always be avoided. We have reviewed two methods for imputing missing data that are similar in many ways (requiring iteration, ending up with a sum of squares) and different in one very important matter. For one, the sum of squares is of a Procrustes distance, and for the other, a bending energy. The choice between the methods is thus, in fact, a choice between these figures of merit. Under

what circumstances would one wish to minimize one of these, or the other?

Thin plate spline interpolation can be considered a reasonable method in the absence of a reference population, as it requires only a single reference form. This single specimen however conveys no information about population variance and covariance; in some cases it may be plausible to match reference and target specimen on some single quantity (for example sex or age).

When specimens are adequate in number, the missing data estimation can be regression-driven instead: The likelihood of the population that includes the resulting landmark configuration is maximized, subject to the proviso that the range of dimensions one intends to impute is actually present in the reference population (e.g. growth allometry).

In the presence of a reference population of which the reconstructed form is likely to be a part, the decision rule we suggest depends on the use to which the completed forms will be put. If you're going to be talking about cranial capacity, distances between landmarks, or other large-scale properties, you want the most precise landmark locations, and so you use the regressions as best you can. If you want to talk about features of shape of the single form, like bumps and bends that might be characters, you use the bending energy.

Take care that the assumptions of the reconstruction must not overlap with the hypothesis you are testing. Statistical reconstruction exploits all information present in the reference population; hence applying any procedure that involves a

covariance matrix (eg. **principal components** or singular warps [Bookstein et al. 2003]), the statistically reconstructed specimen is 'overestimated'. In such cases it makes sense to use geometrical reconstruction instead.

ACKNOWLEDGEMENTS

We want to thank Horst Seidler, Katrin Schäfer, Dennis Slice, Bence Viola, Markus Bernhard and Simon Neubauer for their advice and support. Our research is supported by the AUS-

TRIAN SCIENCE FOUNDATION Project P14738.

¹In practice, given the predictions by each other shape coordinate separately - typically we cannot invert covariance matrices among shape coordinates, so we just add up the separate predictions or the predictions by the first few principal components of those other coordinates.

REFERENCES

- ALLISON, D.P., 2001. Missing Data. Sage University Papers, Thousand Oaks.
- BOOKSTEIN, F., 1991. Morphometric tools for landmark data: geometry and biology. Cambridge University Press, Cambridge (UK)/New York.
- BOOKSTEIN, F., 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med Image Anal* 1:225-43.
- BOOKSTEIN, F.L., GUNZ, P., MITTEROECKER, P., PROSSINGER, H., SCHAEFER, K. and SEIDLER, H., 2003. Cranial integration in Homo: singular warps analysis of the midsagittal plane in ontogeny and evolution. *J Hum Evol* 44(2):167-87.
- DEMPSTER, A.P., LAIRD, N.M. and RUBIN, D.B., 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society* 39:1-38.
- LITTLE, R. and RUBIN, D.B., 2002. Statistical Analysis With Missing Data. John Wiley & Sons, Hoboken, New Jersey.
- MITTEROECKER, P. and GUNZ, P., BOOKSTEIN, F., in press. Semilandmarks in three dimensions. In Slice, D.E. (ed.), *Modern Methods in Anthropology*, Kluwer.
- ROHLF, F.J., 1993. Relative warp analysis and an example of its application to mosquito wings. In Marcus, L.F., Bello, E. and García-Valdecasas, A. (eds.), *Contributions to Morphometrics, Monografías, Museo Nacional de Ciencias Naturales*, Madrid:131-159.
- ROHLF, F.J. and SLICE, D.E., 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst. Zool.* 39:40-59.