# Generalized Procrustes Surface Analysis

Benjamin Pomidor, Dennis Slice
Department of Scientific Computing
Florida State University

Abstract The tools and techniques used in shape analysis have constantly evolved, but their objective remains fixed: to quantify the differences in shape between two objects in a consistent and meaningful manner. The hand-measurements of calipers and protractors of the past have yielded to laser scanners and landmark-placement software, but the process still involves transforming an object's physical shape into a concise set of numerical data that can be readily analyzed by mathematical means.

Here, we present a new method to perform this transformation by taking full advantage of today's high-power computers and high-resolution scanning technology. This method uses surface scans to calculate a shape-difference metric and perform superimposition automatically rather than relying on carefully, tediously, manually placed landmarks. This is accomplished by building upon and extending the Iterative Closest Point algorithm to behave in a manner more conducive to use in morphometric analysis. In particular, we alter the cost function to take both surfaces into account during superimposition so that the superimposition operation is symmetric. We also developed an approach similar to Generalized Procrustes Analysis to handle the superimposition of more than two surfaces at once. We have also examined some ways this new data may be used; we can, for example, calculate an averaged surface directly and visualize point-wise shape information over this surface. Finally, we demonstrate the use of this method on a set of primate skulls and compare the results of the new methodology with traditional geometric morphometric analysis.

## Introduction

- Modern geometric morphometric methods require tediously placed landmarks.
- These landmarks are often taken from surface scans using landmarking software.
- It is difficult to study relatively featureless surfaces with landmark methods.

## Iterative Closest Point (ICP)

- We can circumvent the use of landmarks by instead using the points making up a surface scan.
- The ICP algorithm is commonly used in range image mapping to superimpose surfaces.
- The algorithm works by pairing all the points on one surface with their nearest neighbor, solving a cost function calculated from these pairings to find a minimizing transformation, applying said transformation, and repeating until convergence is reached.

## Symmetry

- Superimposition via the ICP algorithm is not a symmetrical operation.
- In order to calculate a meaningful shape metric, it needs to be symmetrical.
- We do this by redefining the cost function to use the set of nearest neighbor pairings for both surfaces in the cost function simultaneously:

$$C = \sum_{i=1}^{m_A} \left|(p_{A,i} - (Hq_{B,i} - t)) \cdot n_{A,i}\right| + \sum_{j=1}^{m_B} \left|((H^T p_{B,j} + t) - q_{A,j}) \cdot n_{B,j}\right| \qquad (1)$$

- $A$ is the stationary surface, $B$ is the transformed surface
- $m$ is the number of points on a surface
- $q$ is the nearest neighbor of point $p$
- $n$ is the surface normal at point $p$, used to minimize point-to-plane distance
- $H$ and $t$ are rotation and translation terms, which are later applied to $B$

## Prototypes

- The symmetrical ICP formulation works well for superimposing two surfaces, but morphometric data sets typically have tens or hundreds of specimens in a sample.
- By implementing an iterative technique similar to Generalized Procrustes Analysis, we can superimpose all specimens in a data set on a calculated mean surface (the prototype).
- We designate one surface scan as a prototype, superimpose all $n$ specimens in the sample onto the prototype using symmetric ICP, use the $n$ sets of nearest neighbor pairings for the prototype surface to calculate new mean positions for each point in the prototype surface, then repeat the process using the new prototype until the prototype is stable (typically 3-5 iterations).

## Procrustes Surface Metric

- We cannot use the traditional Procrustes distance to calculate inter-specimen shape distances because the points on a surface have no inherent homology and there are different numbers of points per surface.
- With that in mind, we can define the Procrustes Surface Metric, based on root mean square distance:

$$D = \sqrt{\frac{1}{2m_A}\sum_{i=1}^{m_A}(p_{A,i} - q_{B,i})^2 + \frac{1}{2m_B}\sum_{j=1}^{m_B}(p_{B,j} - q_{A,j})^2} \qquad (2)$$

- This metric is very similar to Procrustes distance, and in fact, when calculating the Procrustes Surface Metric between two sets of landmarks, $m_A = m_B$, $p_{A,i} = q_{A,j}$ and $p_{B,j} = q_{B,i}$, and the metric is simply:

$$D = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(p_{A,i} - q_{B,i})^2} \qquad (3)$$

$$D = \sqrt{\frac{1}{m}} \cdot \text{Procrustes distance} \qquad (4)$$

## Data

- The new method was tested using two sets of data (courtesy of David Strait and Jana Makedonska): one with 15 adult male Capuchin monkey skulls and the other with 18 adult male monkey skulls of assorted small primates.
- All skulls had 300,000 - 1,000,000 points in their surface scan as well as 43 manually placed landmarks.
- Both types of data for both data sets were adjusted to a unit centroid size for analysis. Size was later restored to investigate the effect of size on the measurement of form using GPSA.

## Tests and Results

- GPSA was tested comparatively with GPA. All sets of data were superimposed with both methods, then inter-specimen distance were calculated using both Procrustes distance and the Procrustes Surface Metric.
- In the graphs below, from left to right, the data sets used were the single species with size removed, single species with size restored, multiple species with size removed, and multiple species with size restored:

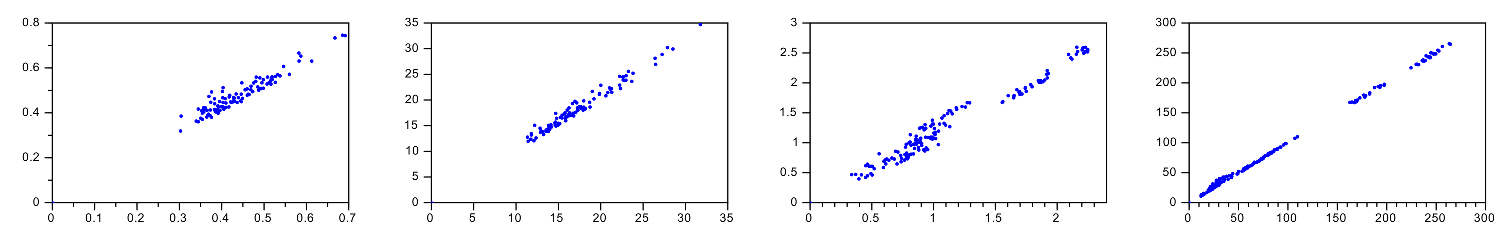

Figure 1: GPSA superimposition vs GPA superimposition measured with Procrustes distance

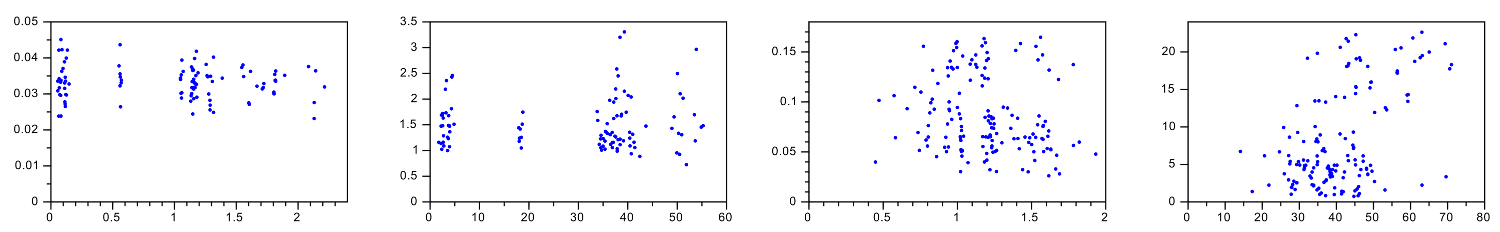

Figure 2: GPSA superimposition vs GPA superimposition measured with Procrustes Surface Metric

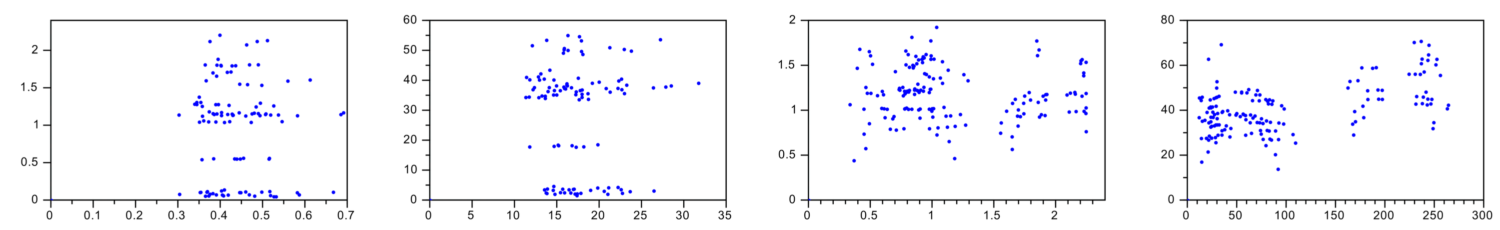

Figure 3: Procrustes Surface Metric vs Procrustes distance using GPA superimposition

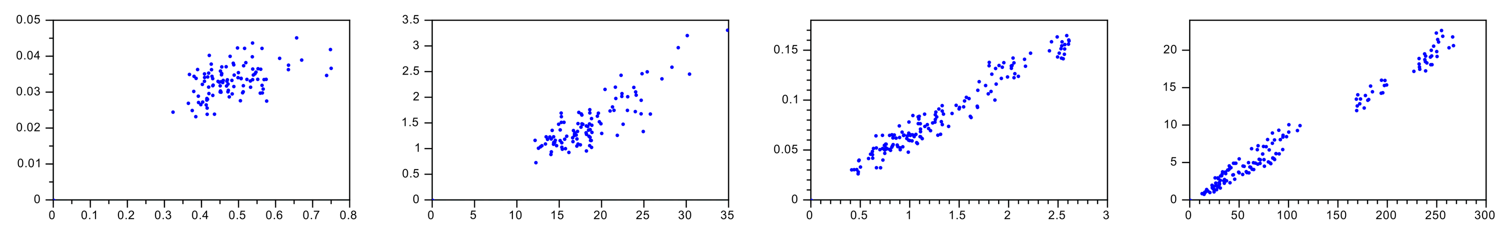

Figure 4: Procrustes Surface Metric vs Procrustes distance using GPSA superimposition

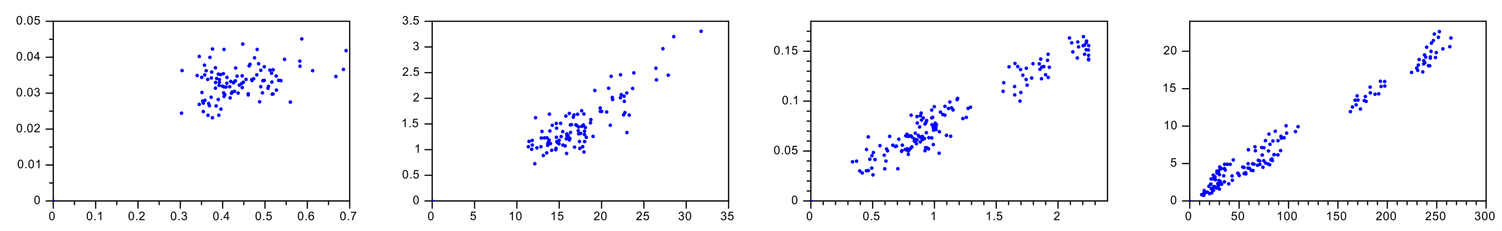

Figure 5: GPSA with Procrustes Surface Metric vs GPA with Procrustes Distance

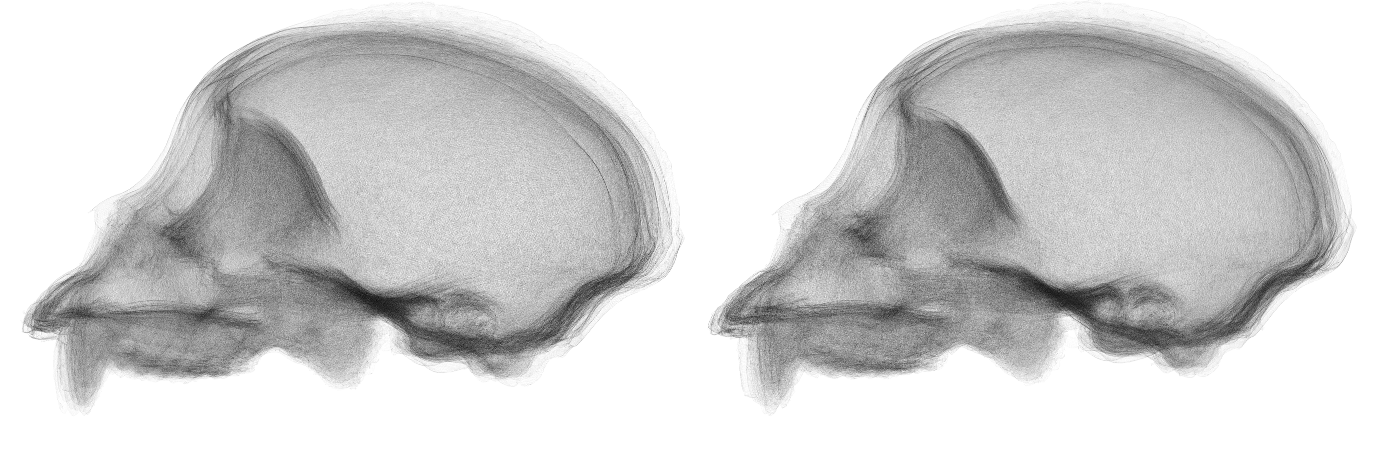

Figure 6: GPA (left) and GPSA (right) superimpositions. Note the slight difference in the vault region.

## Conclusions and Discussion

- In practice (figure 5), GPSA is moderately correlated with GPA, depending on the amount of shape variation in the sample.
- While the superimpositions from GPSA and GPA are similar, they are not identical (figure 6), partially due to the uneven distribution of landmarks. The Procrustes Surface Metric is more sensitive to this difference than Procrustes distance (figures 1 and 2).
- The banded pattern in figures 2 and 3 suggest that there is shape information contained in the surface scans, which is captured by the Procrustes Surface Metric, that isn't represented by landmarks or captured by Procrustes distance.
- The disappearance of the banded pattern in the GPSA-superimposed test (figure 4) indicates that GPSA successfully minimizes this shape difference.
- We may think of landmarks are a low-resolution, unevenly-distributed approximation of shape, while surface scans are very evenly-distributed, high-resolution approximation of shape. The trade-off, as with most higher resolution techniques, is a higher sensitivity to noise, such as differences in occluded regions or incomplete scans.
- Another useful aspect of GPSA is the ability to calculate point-wise shape statistics across the entire prototype surface. The background images are heat maps of the average distance (top) and variance (bottom) for each point on the final prototype surface for the single species data set with size restored.

## References

- Y. Chen and G. Medioni. "Object Modeling by Registration of Multiple Range Images," 1991.
- Guy Godin, Marc Rioux, and Rejean Baribeau. "Three-Dimensional Registration Using Range and Intensity Information,"1994.
- J. C. Gower. "Generalized Procrustes Analysis," 1975.
- Roberto Toldo, Alberto Beinat, and Fabio Crosilla. " Global Registration of Multiple Point Clouds Embedding the Generalized Procrustes Analysis into an ICP Framework," 2010.