

IMP: Superposer

Integrated Morphometrics Package Superposition Tool

Programming by:

David Liebner ('03)
Canisius College
2001 Main St.
Buffalo, NY 14208

Concept by:

H. David Sheets, Dept. of Physics, Canisius College
sheets@canisius.edu

Introduction

Superposer is a part of a suite of software (IMP) meant for use in morphometric analysis. Superposer is intended primarily as a tool for transforming raw morphometric data into one of four superposition types: (1) Bookstein registration, (2) Procrustes superposition, (3) Sliding Baseline Registration, or (4) RFTRA.

How to Use Superposer:

Loading Data

This program is designed to load data files in the x1y1x2y2...CS (*.txt) format or TPS (*.tps) format. If centroid sizes have not been appended to the data set, the program will automatically calculate them and append them to the data.

To load a file, select **File** on the menu bar and click on **Load Data Set**. Choose the file you would like to load. If the file is in the appropriate format, the data will be plotted on the axes, and the filename will appear both in the listbox entitled "Available Data Sets" and underneath the graphical image. The blue asterisks on the graph represent the mean shape for the loaded data.

More than one data set can be loaded in the program at a time. Simply load additional data sets following the same format as above. The filenames for the loaded data sets will appear in the listbox entitled "Available Data Sets." Please note that the active data set is the one that is highlighted in the listbox, *not* necessarily the one that is displayed on the graph. To switch between active data sets, click on the filename of the desired data set to highlight it. This will allow the program to carry out superpositions with that particular data set.

Carrying out a Superposition

Once the appropriate data set has been loaded and selected, a superposition can be carried out. A number of different options are offered:

1) *Select Superposition Type*

Select from among four different superposition types: Procrustes, Bookstein, Sliding Baseline, or RFTRA. Depending on which superposition type is chosen, it will limit which of the remaining options can be selected.

Procrustes Superposition: Individual shapes in the data set are superimposed on a reference shape. The superposition is optimized by minimizing an appropriately chosen error function that incorporates the distances between landmarks on the image and their respective landmarks on the superimposed shape. Images may be constrained to be centroid size and/or centered about the origin.

Bookstein Registration: Two landmarks are selected to serve as anchoring points for the image. One landmark is fixed at the origin (0,0). The second landmark is fixed at the point (1,0).

Sliding Baseline Registration: Once again, two landmarks serve as anchors for the image. Here, the two landmarks are constrained to lie along the x-axis. The image is often constrained to be centroid size and centered about the y-axis. If it is not fully constrained in this manner, a Procrustes-type error function may be minimized to optimize the fit. Sliding baseline registration is especially useful when the images being superimposed have a well-defined axis of symmetry.

RFTRA: This is a standardized robust superposition. Initially, the shape data is superimposed using a least-squares Procrustes superposition in which the shapes are constrained to be centroid size and centered about the origin. Subsequently, a series of repeated medians is carried out to determine the median coefficients for scaling, rotation, and translation. The result is a robust superposition that is resistant to outlying data points. One disadvantage to the technique, however, is that the final superposition is somewhat dependent on the initial orientation of the reference object.

2) *Select Landmarks that Form the Baseline* (Bookstein and Sliding Baseline, only)

Enter the numbers of the landmarks that will be the anchors for the superposition. Landmark numbers are based on the order that each landmark is entered in the data set. (If the data file that is going to be used in the

superposition is currently graphed on the axes, see “View Options” to see how landmark numbers can be displayed on the graph.)

3) *Select Error Function* (Procrustes and Sliding Baseline, only)

A variety of error functions are offered for Procrustes and Sliding Baseline superpositions. Most error functions are constructed using the following standard format:

$$Error = \sum_{i=1}^n \sum_{j=1}^k f(d_{ij})$$

Here, n denotes the number of distinct images stored in the data file, while k denotes the number of landmarks contained within each image. The term d_{ij} is used to denote the geometric distance between landmark j on the i^{th} object and landmark j on the reference object. The error function f is subject to user specification. Currently, there are five error functions available to choose from. Least-squares is standard for most purposes; however, it may not be ideal when variations in shape are highly localized, or when there are outliers. In these cases, one should consider one of the four other error functions, which are more robust in these situations.

Error Function	Format
Least squares	$Error = \sum_{i=1}^n \sum_{j=1}^k d_{ij}^2$
Absolute value	$Error = \sum_{i=1}^n \sum_{j=1}^k d_{ij}$
Andrew’s sine	$Error = \sum_{i=1}^n \sum_{j=1}^k \sin(d_{ij})$
Natural logarithm	$Error = \sum_{i=1}^n \sum_{j=1}^k \ln(d_{ij} + 1)$
Least median	$Error = \sum_{i=1}^n \text{median}(\{d_{i1}, d_{i2}, \dots, d_{ik}\})$

It should also be noted that the “mean” that is calculated during the superposition routine for both Procrustes and Sliding Baseline is dependent on the choice of error function. Following the completion of the superposition, the “mean” is calculated by minimizing the selected error function with d_{ij} now used to represent the distance between landmark j on the i^{th} object and landmark j on the “mean.” The “mean” is also constrained to be both centroid size and centered. (This technique is used in order to force the “mean” to lie within the same dimensional space as the superimposed images. Also, it is important to note that if the true mean were used instead, it would be

impossible to carry out a superposition around that mean using nonstandard error functions.)

4) *Select a Reference Image* (Procrustes, Sliding Baseline, and RFTRA)

For the above superposition types, a reference image is required to carry out the superposition. Two options are presented:

Superimpose objects around “mean” of current data set: As indicated above, the “mean” is not a true mean, but it *is* a composite of the shapes in the data set as defined by the selected error function. Superpositions around the mean are often used to show the extent of variance relative to certain landmarks.

Superimpose objects around reference form: In this case, the superposition is carried out around an image that is separately loaded into the program. This option cannot be selected unless a reference file has already been loaded. To load a reference form, select **File** on the menu bar and click on **Load Reference Form**. Choose the appropriate file (*.txt or *.tps format). Make sure that the file contains only a single shape image; otherwise, it will not be loaded. Once the file has been loaded, the title of the file should appear in the popup box that is adjacent to the radio button labeled “Reference form.” More than one reference form can be stored in the program at a time. Simply repeat the above procedure. A superposition will be carried out around a reference form only if the appropriate radio button is highlighted, and only if the reference form is displayed in the popup box.

5) *Additional Options* (Procrustes and Sliding Baseline, only)

Two additional options are presented for Procrustes and Sliding Baseline superpositions.

Align all centroids: For a Procrustes superposition, this forces the images to be centered at the origin. For a Sliding Baseline registration, the images are centered around the y-axis.

Set centroid size to unity: All images are forced to be unit centroid size.

For standard least squares superpositions, it is faster and more convenient to select both of these options. For nonstandard error functions, the superposition may be more optimal if one or both of these options are not selected, since this allows for greater freedom with outliers.

Once all of the options have been selected, the superposition is ready to be carried out. Simply click on the pushbutton labeled, “Do it!” to start the superposition. The pointer should change to a timepiece to indicate that the software is working. When the superposition is finished, the program will display the finished superposition on the

graph. The specifics of the superposition will be listed below the graph, including (as relevant) the filename, whether an independent reference form was used, landmarks used in the baseline, error function selected, time required to carry out the superposition, average error per sample, and any other options.

Once a superposition has been carried out, it is stored in the computer's RAM. To retrieve the results of a superposition that was carried out earlier in a session, simply select the same options as before, and click on "Do it!" The results should be rapidly pulled up.

View Options

A number of different viewing options are presented for the data under the **View** menu:

- 1) **Autoscale, True Aspect:** This option is selected by default. This forces the axes on the graphical image to be scaled in equal increments. If this option is unchecked, the graphical image is displayed in the default format—i.e. the axes tightly surround the image on all sides.
- 2) **Show Data Labels:** This option, by default, is unchecked. When it is selected, the program will automatically display the landmark numbers next to the appropriate data points for the graphical object that is currently displayed on the graph.
- 3) **Hide Axes:** When this option is selected, the axes background and tick marks are not displayed. Only the graphical image appears.
- 4) **Only Show Mean:** This option was designed to allow only the average position of the data set to be displayed. The individual data points themselves are hidden.
- 5) **Show Reference:** If an independent reference form was selected for the superposition, it can be displayed on the graph. Simply select this option. The reference form will appear with green asterisks. If no reference form was used in the superposition, this option is not active.

These options are often most useful when copying or printing images in order to tailor their appearance.

Copying Images

In order to copy an image to the clipboard, click on **Edit**, then **Copy to Clipboard**. The image will be copied as a meta-file.

In order to copy an image to a file, click on **Edit**, then **Copy to File**. Select an appropriate filename and click **OK**. The image will be copied as an Encapsulated Post-

Script (*.eps) file. These images should load normally into most word-processing programs.

Printing Images

To print the image that is displayed on the screen, simply click on **File** and select **Print**. Click **OK** when ready to print.

Saving Data

Two save options are presented. In order to save the superpositioned images that are displayed *on the graph*, click on **File** and select **Save Current Data**. Choose an appropriate filename, and click on **OK**. This will save the superimposed data *only*. It will *not* save the reference or the calculated mean. All data is saved in the x1y1x2y2...CS format.

In order to save the calculated mean for the data displayed *on the graph*, click on **File** and select **Save Current Mean**. Once again, choose an appropriate filename, and click on **OK**. This will save the calculated mean *only*. Once again, the mean is saved in the x1y1x2y2...CS format.

Conclusion

Superposer is a fairly versatile tool in that it offers a variety of superposition options, both standard and robust. Hopefully, it can help you in your endeavors at shape analysis.