A computational method for age-at-death estimation based on the pubic symphysis

Detelina Stoyanova, Dennis Slice

Department of Scientific Computing
Florida State University

A significant component of forensic science is analyzing bones to assess the age-at-death of an individual. Forensic anthropologists often include the pubic symphysis in their studies. Subjective methods, such as the Suchey-Brooks method, are currently used to analyze the pubic symphysis. This study examines a more objective, quantitative method. The method analyzes 3D surface scans of the pubic symphysis and implements a thin plate spline algorithm which models the bending of a flat plane to approximately match the surface of the bone. The algorithm minimizes the bending energy required for this transformation. Results presented here show that there is a correlation between the minimum bending energy and the age-at-death of the individual. The method could be useful to medical practitioners.

Introduction

- The pubic symphysis is a small surface connecting the left and right pelvic bones in the front of the pelvic girdle.
- Throughout life, the surface of the pubic symphysis changes at a more-or-less predictable rate.
- For many decades, researchers have been interested in developing techniques for age estimation based on the changing morphology of the pubic symphysis. Such methods include Todd, McKern and Stewart and the most popular one - the Suchey-Brooks method.
- There is a definite need for an objective, quantitative method.

Method Overview

- The method uses 41 scans from white male individuals with known ages.
- It reads in and cleans scan data from ASCII PLY files.
- It applies Principal Component Analysis (PCA) to standardize the position of the scans in space.
- It selects two sets of equidistant control points forming square grids.
- It uses thin plate splines (TPS) to give each scan surface a quantitative representation.

3D-scan data

The purpose of 3D scanners is to generate a dense point cloud or a polygonal mesh. They capture the geometry of a physical object with hundreds or thousands of measurements. The 3D shape is represented as numerous small adjacent triangles that are called facets. Every face has three vertices and each vertex has three coordinates that specify its location in space. The data from the scans are stored in binary or text files that are essentially lists of numbers. The files store the coordinates of all the vertices and the information on how they are connected to form the triangles. Examples of 3D surface scans of pubic symphysis are shown below.

Principal Component Analysis (PCA)

The next step of our method is standardizing the orientation of the scans. It starts by translating the scan to the center of the coordinate system and is followed by PCA. The goal is to position the bone in such a way that its center matches the center of the coordinate system and the x, y, and z axes define the dimensions with the largest variance (x), the second largest variance (y) and the smallest variance (z). This way the x − y plane approximates the articular surface of the bone, and we can study the distances in the z direction.

Algorithm

In the current application, the PCA algorithm uses a mean-centered data matrix, A, that is an n × 3 matrix where n is the number of vertices and the three columns correspond to the x, y and z coordinates of each vertex after the mean-centering. First, the 3 × 3 matrix defined by \( A^T A \) is calculated. To do an eigenanalysis on the matrix we use singular value decomposition. The \( A^T A \) matrix is factored into a 3 × 3-diagonal matrix \( S \) that has the singular values as its diagonal entries and two orthogonal 3 × 3 matrices \( U \) and \( V \) (in our case \( U = V \) because \( A^T A \) is a symmetric matrix) such that

\[
A^T A = USV^T.
\]

The n × 3 matrix \( AV \) gives the coordinates for the vertices after the PCA rotation.

Selecting control points

To select the first set of control points (the points lying on the flat plane), the method scales the bones to be of uniform length equal (in our case) to 20. It then selects 51 equidistant control points in the x-direction and however many points fit in the y-direction. This results in a square uniform grid of size 80 × 80. We can choose any length and number of points but those numbers are used for convenience. Then, we find the projection of the first set of control points onto a flat triangle (i.e. the scan). The projection points become the second set of control points. A non-existent projection means that the given point is outside of the surface scan and is removed.

Thin Plate Splines

Given the two sets of control points, one lying in a plane and the other consisting of corresponding points on the symphysis surface, we can bend the plane so that the two sets of points match exactly.

The implementation of the method involves solving a linear system \( \mathbf{Ax} = \mathbf{b} \), and finding a solution vector \( \mathbf{x} \). The components of the solution vector may be used to interpolate the rest of the points to find the minimum bending energy required for bending the plane to match the surface of the pubic symphysis. Let \( \mathbf{x} \) be the first \( k \) components of \( \mathbf{x} \) where \( k \) is the number of control points in one of the 2 sets and \( a_1, a_2, a_3 \) the last 3 elements of \( \mathbf{u} \). To interpolate the points we can use the exact equation for \( \mathbf{a} \) that is

\[
\mathbf{a} = \mathbf{a}_1 + a_2 \mathbf{v}_2 + a_3 \mathbf{v}_3 + \sum_{i=1}^{n} \alpha_i [\mathbf{plane} - \mathbf{plane}(\mathbf{a})].
\]

The value of the minimum bending energy is given by \( E_{min} = \mathbf{a}^T \mathbf{C} \mathbf{a} \). The minimum bending energy is a measure of the complexity of the surface scan, which is, in part, the basis of the subjective methods currently in use. A correlation between the age-at-death and the minimum bending energy could be used to estimate the age-at-death.

Future Work

- Outlier Analysis.
- Different methods for analyzing the surface.
- Using two different regressions - one for younger people and one for older people.
- and more ...

References


Results

The results presented here were generated using 51 control points in one set.

<table>
<thead>
<tr>
<th>Control Point</th>
<th>Linear Regression</th>
<th>Coefficient</th>
<th>Mean</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.4914</td>
<td>0.0004</td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>RMSE</td>
<td>17.5899</td>
<td>0.8524</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A plot of the results showing log(age-at-death) vs log(bending energy) and the best fitting line that goes through the points.

Position with axis after PCA
Uniform square grid
First set of control points (on the plane)