Statistical 3D Cranio-Facial Models

Maxime Berar
Laboratoire des images et des signaux, Saint Martin d'Hères, France
berar@lis.inpg.fr

Michel Desvignes
Laboratoire des images et des signaux, Saint Martin d'Hères, France
Michel.Desvignes @lis.inpg.fr

Gérard Bailly
Institut de la Communication Parlée, Grenoble, France
Gérard.Bailly @icp.inpg.fr

Yohan Payan
Laboratoire TIMC/IMAG, Faculté de Médecine, La Tronche France
Yohan.Payan @imag.fr

Abstract

In forensic science, 3D cranio facail reconstruction is used to reconstruct the face from a skull. This can be done by manual approaches or computer assisted methods. The proposed statistical model represents the relationship between the skull and the soft tissues and is inverted to reconstruct the unknown face from the known skull. It is a specific application of the missing or occulted data problem. Results are visually correct.

1. Introduction

In forensic science, face reconstruction try to estimate the face of an unknown human, from his skull. Manual methods [1] use clay or platicine with a very impressive visual quality. The main shortcomings are subjectivity and time consuming. Technology, 3D graphics and virtual reality have introduced computer assisted methods, which can reduced this time.

Some current computerized techniques either fit a template skin surface to a set of interactively placed virtual dowels on a 3D digitized model of the remaining skull [2]. Other ones propose to deform a reference skull in order to match the remaining skull, thanks to crest lines (lines of maximal local curvature) [3] or feature points [4]. Then they apply an extrapolation of the calculated skull deformation to the template skin surface. Recent works using a combined statistical deformable model of facial surfaces and tissue thickness [5] addresses the facial reconstruction problem.

However, this problem can be seen as a missing data or pattern recognition problem : if a model of the face and skull exists, an occulted or missing part (the face), must be determined from a known part (the skull). Statistical pattern recognition techniques solve this problem in an elegant way.

2. Statistical Model

2.1. Building a 3D face and skull statistical model

As in computer graphics and visualization, an object is represented as a mesh, a set of vertices connected by arcs. Each vertex of the mesh is supposed to be a semi-landmark of the 3D surfaces, i.e. a point with no anatomic signification but which is the same for everybody. A statistical model of the (partial) skull and the face is then built using Principal Component Analysis (PCA). The result of the PCA is a geometrically averaged facial template, which is computed together with a correlation-ranked set of modes of principal variations based on inter-subjects variations.

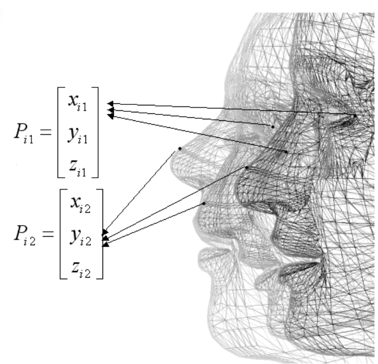

Figure 1. Each vertex of the subject-specific meshes is considered as the same location reflecting thus the inter-individual variations of shape

Proceedings of The Sixth IEEE International Conference on Computer and Information Technology (CIT'06)
0-7695-2687-X/06 $20.00 © 2006 IEEE

Let $X_i = (x_{i1}, y_{i1}, z_{i1}, \ldots, x_{in}, y_{in}, z_{in}) \in R^{3n}$ be the locations of n vertices of the normalized meshes. Using PCA, we can write $T \approx \tilde{T} + \Phi b$, where is the mean mesh of the skull and face, $\Phi = (\phi_1 \mid \ldots \mid \phi_t)$ is a $(n+m) \times (n+m)$ matrix composed with the eigenvectors of the covariance matrix S of the centered data and b is a vector of t dimension : $b = \Phi^t (T_i - \overline{T})$.

The dimension t of the vector b is the number of eigenvectors with the largest eigenvalues. In classical use of PCA, such as de-noising, t is chosen by $\sum_{i=1}^{t} \lambda_i \geq 0.95 \sum_{i=1}^{m+n} \lambda_i$. The vector b is then a good approximation for the original dataset and any of $n+m$ points can represented or retrieved with the $t_{t<n+m}$ values of the vector b by $T \approx \tilde{T} + \Phi b$

2.2. Missing Data and Recognition

The linear PCA model defined here can be extended in an elegant way in order to take into account spatial relations between landmarks. This model can be used to estimate an unknown part of a partially visible or occulted model [6].

Under this hypothesis, if some points (says $t=n$ points) are known, the remaining unknown points are determined using PCA. Without any approximations, we can write:

$$\begin{bmatrix} C_1 \\ \vdots \\ C_n \\ X_1 \\ \vdots \\ X_m \end{bmatrix} = \begin{bmatrix} \overline{C}_1 \\ \vdots \\ \overline{C}_n \\ \overline{X}_1 \\ \vdots \\ \overline{X}_m \end{bmatrix} + \begin{bmatrix} \Phi_{1,1} & \cdots & \Phi_{1,n+m} \\ \vdots & \ddots & \vdots \\ \Phi_{n+m,1} & \cdots & \Phi_{n+m,n+m} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \\ \beta_{n+1} \\ \vdots \\ \beta_{n+m} \end{bmatrix}$$

This is a linear system with $n+m$ equations and unknowns that can not be resolved. Since PCA can represent the dataset with $t<n+m$ values, suppose $t=n$, the unknown vector $(\beta_1, \ldots, \beta_n, X_1, \ldots, X_m)$ in the following system. If t<n, the system becomes overdetermined and can be written as :

$$\begin{bmatrix} C \\ 0 \end{bmatrix} = \begin{bmatrix} A & 0 \\ B & -Id \end{bmatrix} \begin{bmatrix} \beta \\ X \end{bmatrix} \; or \; Y = MZ$$

$$where \; A = \begin{bmatrix} \Phi_{1,1} & \cdots & \Phi_{1,j} \\ \vdots & \ddots & \vdots \\ \Phi_{n,1} & \cdots & \Phi_{n,j} \end{bmatrix} and \; B = \begin{bmatrix} \Phi_{n+1,1} & \cdots & \Phi_{n+1,j} \\ \vdots & \ddots & \vdots \\ \Phi_{n+m,1} & \cdots & \Phi_{n+m,j} \end{bmatrix}$$

2.2.1. First solution

Now, with this particular form, the system is easily solved by using the first part of the system, as in the classical regression model:

$$C = A * \beta \text{ i.e. } \beta = A^{-1} * C$$

and then

$$X = B * A^{-1} * C$$

2.2.2. Second solution

Another way is to use a least square method to solve the system. The cost function is

$$J(\beta, X) = \begin{bmatrix} C - A*\beta & -X + B*\beta \end{bmatrix} \begin{bmatrix} C - A*\beta \\ -X + B*\beta \end{bmatrix}$$

In matricial form, J is rewrite :

$$J(Z) = J(\beta, X) = \left\| Y - M*Z \right\|^2 =$$

$$\left(Y'Y - Y'MZ - Z'M'Y + Z'M'MZ \right) = (Y'Y - 2Y'MZ + Z'M'MZ)$$

As the matrix [A B] is an orthonormal basis, M'M takes the following form and the cost function becomes:

$$M'M = \begin{bmatrix} A' & B' \\ 0 & -Id \end{bmatrix} \begin{bmatrix} A & 0 \\ B & -Id \end{bmatrix} = \begin{bmatrix} Id & -B' \\ -B & Id \end{bmatrix}$$

$$J(\beta, X) = C'C - 2\beta' A'C + X'X - 2\beta' B'X + \beta'\beta$$

The derivatives with respect to X and β are null :

$$\frac{\partial J(\beta, X)}{\partial \beta} = -2A'C + 2\beta - 2B'X = 0 \Rightarrow \beta = A'C + B'X$$

$$\frac{\partial J(\beta, X)}{\partial X} = -2B\beta + 2X = 0 \Rightarrow X = B\beta$$

Reporting the second equation in the first one, the solution is :

$$\beta = \left(Id - B'B \right)^{-1} A'C$$

$$X = B\left(Id - B'B \right)^{-1} A'C$$

Note that (Id-B'B) is always invertible, since it is a symmetric positive defined matrix.

In this framework, a linear approximation of spatial relations between known and unknown points is explicitly determined from the eigenvectors of the covariance matrix.

2.3 Results on Synthetic data

These two methods have been applied on synthetic data sets. A set of 200 examples 1000 points with known linear relationship have been generated. Then, a variable part of them have been suppressed and we have computed the error of the reconstruction of this missing part.

We have also evaluated the difference between the 2 presented methods. On figure 2, the values of parameters are drawn with respect to the number t of variability parameters used for reconstruction. The second method, black dot line is always at least 10% better than the first one drawn in blue, plain line.

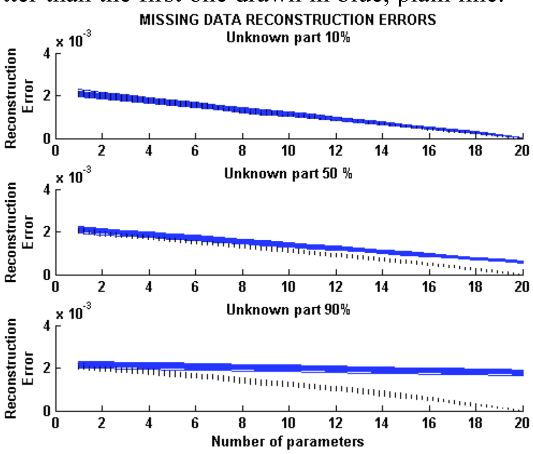

Figure 2. Reconstruction errors for variable unknown parts of the data for the first method (blue plain line) and for the second method (black dot line).

3. Application to 3D facial reconstruction

3.1. General approach

The general scheme of our approach is presented on figure 3.

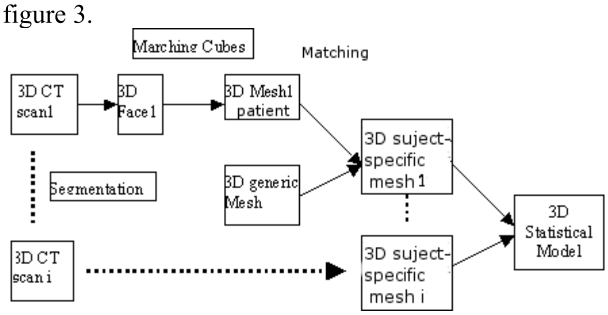

Figure 3. Building a 3D statistical model from 3D CT Scans.

Coronal CT slices (see Figure 4.) were collected for the partial skulls and faces of 15 subjects (helical scan with a 1-mm pitch and slices reconstructed every 0.31 mm or 0.48 mm).

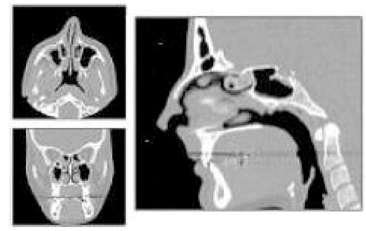

Figure 4. 3D raw scan data.

Segmentation is the first step to get 3 binary volumes (skull, face, jaw) from 3D CT scans. A threshold is easily defined for face object since CT exhibits a large contrast between soft tissues and air. The skull is obtained from a pile of thresholded 2D images. A 3D binary image is reconstructed after 2D hole filling. Some CT artifacts (filling of teeth) are manually discarded

In a second step, three meshes are obtained from the marching cubes algorithm. Patient meshes for the skull, jaw and face have around 180000, 30000 and 22000 vertices.

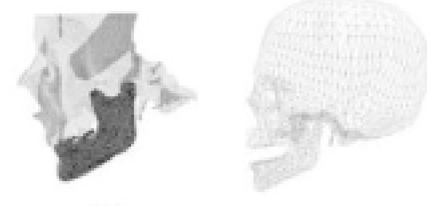

Figure 5. (a) shape reconstructed using the marching cube algorithm; (b) generic mesh obtained from the Visible Woman Project®

The assumption of semi landmarks is assumed by a generic mesh, with a fixed number of vertices. In the third step, this generic mesh is registered on each patient's mesh, which becomes a 3D subject-specific mesh. These meshes have the same number of vertices for every patient (3473 for the skull, 1100 for the jaw 5828 for the face). and a vertex matches the same point for every patient. A 3D-to-3D meshes matching algorithm is used to obtain these normalized meshes.

A vector is formed form the coordinates of the ordered vertices. in the patient specific mesh. The statistical model is then built from these vectors. In the facial reconstruction, the skull and jaw are known and we try to determine the face from the model. Then, this is a missing data problem and the face is determined by the previously presented method.

3.2 Results

An example of this statistical reconstruction is presented figure 6. As we can show, the global aspect is coherent, and the reconstruction is correct.

0-7695-2687-X/06 $20.00 © 2006 IEEE

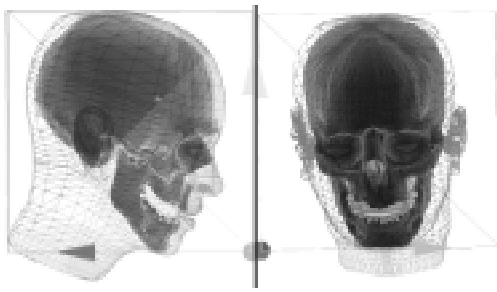

Figure 6. Example of facial reconstruction

To quantify the results, we compute the error between the true value of the coordinates and the computed value.

As the number of samples of CT scans is low, we used a leave-one-out approach to quantify the accuracy of the facial reconstruction. One CT scans is discard from the learning database and becomes the test sample. Every subject becomes the test sample in turn.

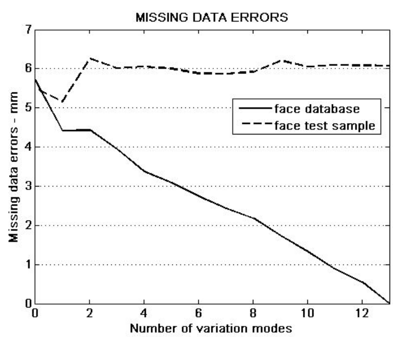

Figure 7. Reconstruction error of the face.

Figure 7 gives the mean reconstruction error versus the number of variability parameters, for the samples of the learning database in plain line. When computed with the complete variation modes, the reconstructions of the learned sample have an accuracy of 0,5mm.

For the test sample, the reconstruction error is about 5mm in the better case. This result is explained by the low number of sample in the training set, as the error on training set is low. With this low number, the variance-covariance computed to build the model have a low accuracy. The relationships between points cannot be generalized with a good confidence.

3.3. Synthetic facial database

As the number of CT scans is limited, we have built a larger synthetic training database. One face and skull mesh has been chosen and a set of elastic transforms is applied to this mesh. A gaussian noise is added to each vertex oh the mesh. The training database is generated with one hundred meshes.

In this experiment, the accuracy of the reconstruction of the face from the skull is about 1mm (figure 8) for a gaussian noise of 2mm.

With a gaussian noise of 5mm, the accuracy of the face from the skull is 5mm, as with the real data.

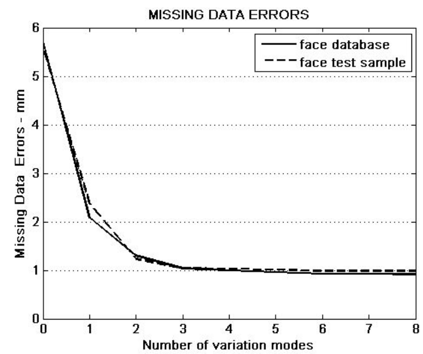

Figure 8. Reconstruction error of the synthetic face

Conclusion

In this paper, we have developed a statistical model for pattern recognition. With this model, we have defined a method to determine missing or occulted part of an object. It has been applied to the 3D facial reconstruction problem, where the face must be determined from the skull. He results are perceptually correct and quantitative measurements show that the training set must be larger.

References

[1] C. Wilkinson, *Forensic Facial Reconstruction*: Cambridge University Press, 2004

[2] J. G. Clement, M. K. Marks: *Computer-Graphic Facial Reconstruction*, Academic Press, 2005

[3] G. Quatrehomme et al, "A fully three-dimensional method for facial reconstruction based on deformable models", *Journal of Forensic Science*, AAFS, 1997, pp. 649-652.

[4] M.W. Jones, "Facial reconstruction using volumetric data.", IVMVC, Germany , 2001, pp. 21-23.

[5] P. Claes, D. Vandermeulen et al., "Statistically deformable face models for cranio-facial reconstruction", ISPA, Croatia, 2005, pp. 347-352.

[6] M. Berar, M. Desvignes et al. Missing data estimation using polynomial kernel, LNCS 3686, UK, 2005,pp 390-399.