Face Reconstruction from Skull Shapes and Physical Attributes

Pascal Paysan[1], Marcel Lüthi[1], Thomas Albrecht[1], Anita Lerch[1], Brian Amberg[1], Francesco Santini[2], and Thomas Vetter[1]

[1] Computer Science Department, University of Basel, Switzerland
[2]Division of Radiological Physics, University of Basel Hospital, Switzerland
{marcel.luethi, thomas.albrecht, anita.lerch, brian.amberg, francesco.santini, thomas.vetter}@unibas.ch

Abstract. Reconstructing a person's face from its skeletal remains is a task that has over many decades fascinated artist and scientist alike. In this paper we treat facial reconstruction as a machine learning problem. We use separate statistical shape models to represent the skull and face morphology. We learn the relationship between the parameters of the models by fitting them to a set of MR images of the head and using ridge regression on the resulting model parameters. Since the facial shape is not uniquely defined by the skull shape, we allow to specify target attributes, such as age or weight. Our experiments show that the reconstruction results are generally close to the original face, and that by specifying the right attributes the perceptual and measured difference between the original and the predicted face is reduced.

1 Introduction

Face reconstruction from skeletal remains has been practiced for well over a hundred years and is now an important technique in forensic science. Apart from its practical application, facial reconstruction also makes a great machine learning task. Given a set of training images depicting both the face and skull, can we learn a mapping from these data sets which predicts the correct face surface for a given skull?

In this paper we propose to model the normal facial surface and skull morphology by means of two separate statistical shape models. We use a shape fitting algorithm to fit the statistical models to Magnetic Resonance (MR) images of the human head. Face reconstruction becomes the problem of learning the relation between the skull and face model parameters. More generally, our method can be seen as an attempt to learn the relationship between two separately constructed but dependent shape models. This makes it possible to use the statistical information represented in one model when given an observation for the other model.

In the field of facial reconstruction, two schools of thoughts have developed [1]: Practitioners of the first school think that all reconstruction methods are inexact and the true face can only be approximated by a facial type which characterizes many possible faces. The second school of thought is dominated by the belief that the facial morphology can be determined from the skull with such accuracy as to make the individual recognizable by including subtle characteristic details of the skull morphology into the analysis. Our method combines features of both schools of thought. The shape models represent both the general shape as well as the typical details of the individual's morphology. It is clear, however, that even when a perfect reconstruction of the facial shape can be achieved, the relationship between skulls and faces is not one-to-one. The face of a single person can change with age or weight while the skull remains the same. Our method therefore allows to constrain the possible reconstructions by specifying such attributes. Our experiments show that correctly specified attributes lead to more accurate reconstructions. Moreover, different reconstruction results for the same individual can be computed, which has been hypothesized to make recognition easier [2].

Related Work While traditional methods of modeling the reconstruction using clay are still in use, many methods for facial reconstruction based on 3D computer graphics have been developed. A recent review of current methods is given in [1]. Early approaches mimicked the manual approach and simply deform a template face to match the typical soft tissue thickness to discrete markers [3, 4] or full soft-tissue maps [5]. These technologies also provide the key to recent methods based on statistical shape models [6–8]. Claes et al. [6] use a statistical face model and incorporate properties such as BMI, age and gender. In contrast to our method, the fitting of the face model is performed by simple interpolation of skin markers. Tu et al. [8] use warping techniques to align skulls from a training set to a new skull. After registration, a PCA model is built from the remaining differences in facial shape. This model captures the variation due to factors such as weight and age. Closest to our work is the work by Berar et al. [7]. They build a joint statistical model of face and skull shape. Face reconstruction is treated as a missing data problem, which has a straight-forward solution. While a similar goal as ours can be achieved, the model can only be built from data sets which clearly show both the skull and the face surface. This is an important difference in practice, as this currently requires the use of CT images, which are much more difficult to obtain.

2 Background

At the core of our method are the statistical shape models. They efficiently capture the shape properties and guarantee that only statistically likely shapes are represented.

2.1 Statistical Shape Models

Statistical shape models are a widely used tool in computer vision and medical imaging. While the method is independent of the kind of shape model used, we use a *Morphable Model*, which is obtained by applying Principal Component Analysis (PCA) to data sets for which dense point-to-point correspondence has been established. From n data sets, represented by vectors $\boldsymbol{s}_i \in \mathbb{R}^m$, the mean $\bar{\boldsymbol{s}}$ and covariance matrix $\boldsymbol{\Sigma}$ are calculated. PCA consists of an eigenvalue decomposition $\boldsymbol{\Sigma} = \boldsymbol{U}\boldsymbol{D}^2\boldsymbol{U}^T$, where $\boldsymbol{U}$ is the orthonormal matrix of the eigenvectors of $\boldsymbol{\Sigma}$, and $\boldsymbol{D}^2$ is a diagonal matrix with the corresponding eigenvalues. With the help of a coefficient vector $\boldsymbol{\alpha}$, each shape can be represented as a linear combination of the eigenvectors:

$$\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{\alpha}) = \boldsymbol{U}\boldsymbol{D}\,\boldsymbol{\alpha} + \bar{\boldsymbol{s}}. \tag{1}$$

When constructing a PCA-based morphable model, it is assumed that the shape vectors $\boldsymbol{s}$ are distributed according to a multivariate normal distribution $\mathcal{N}(\bar{\mathbf{s}}, \boldsymbol{\Sigma})$. Thanks to the representation in Equation 1, the density function takes the simple form:

$$p(\boldsymbol{s}(\boldsymbol{\alpha})) \propto \exp(-\|\boldsymbol{\alpha}\|^2). \tag{2}$$

2.2 Training Data

Three different data sets are used for reconstruction, face scans, skull scans and anchor examples, as illustrated in Figure 1a. The face model we use in our experiments is built from 840 structured light 3D surface scans. For each scanned individual, a number of attributes such as age, weight, and gender were recorded in addition to the geometry and texture of the faces. As most of these attributes can be considered to be independent of the skull shape, we can use them to manipulate the predicted face. We can learn the relationship $\boldsymbol{\vartheta} = f(\boldsymbol{\alpha})$ between the model parameters $\boldsymbol{\alpha}$ and the attributes $\boldsymbol{\vartheta}$ by using a regression method (Figure 1a). For the actual face reconstruction, only the most significant $m_f = 50$ principal components are used.

The skull model consists of $m_s = 20$ segmented CT scans. Its parameters are denoted by $\boldsymbol{\beta}$. It is extremely difficult to obtain CT data sets of the full head of healthy persons, as the scanning process exposes the patients to harmful radiation. Our CT data set therefore includes many scans of dry skulls, which are more easily acquired in sufficient quality.

In order to establish a connection between the face and the skull model, we have acquired a third data set of $n = 23$ MR Images, where both the skull and the face are visible. They can be used as anchor points between the skull and the face model. We can fit both models to these "anchor examples", yielding pairs $(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i)$ of face model parameters $\boldsymbol{\alpha}_i$ and skull model parameters $\boldsymbol{\beta}_i$. [1]

2.3 Statistical Model Fitting

Given an MR image of the head, the goal of model fitting is to find a parameter vector $\boldsymbol{\alpha}$, such that the shape $\boldsymbol{s}(\boldsymbol{\alpha})$ matches the corresponding face or skull contour in the MR image. Moreover, it should be a likely instance of the shape, i.e. we require $\|\boldsymbol{\alpha}\|^2$ to be small (cf. Equation (2)). More formally, let $S \subset \mathbb{R}^3$ be the contour in the image and let $D_S[S']$ be a function measuring the distance between the contour S and S'. The optimal parameters are given as the solution to the optimization problem:

$$\min_{s,t,R,\alpha} D_S[s\boldsymbol{R}(\bar{\mathbf{s}} + \boldsymbol{U}\boldsymbol{\alpha}) + \boldsymbol{t}] + \lambda\|\boldsymbol{\alpha}\|^2 \tag{3}$$

where $s \in \mathbb{R}$ is a scaling factor, $\boldsymbol{t} \in \mathbb{R}^3$ a translation, $\boldsymbol{R} \in \mathbb{R}^{3\times 3}$ a rotation matrix and λ a weighting coefficient. For more details we refer the reader to [9].

[1] It is not possible to use the MR images directly to build the skull model, as skull segmentation from MR images requires a strong shape prior (which is, in our case, the (CT) skull model).

3 Method

As discussed above, the relationship between skulls and faces is not one-to-one. The face shapes offer much more flexibility than the skulls. In our case, this effect is amplified by the number of training examples for the face model being much larger than the number of skulls and anchor examples. We take advantage of this additional flexibility by reconstructing the face not only from the skull but from the skull *and* a set of attributes. In this way, we can reconstruct faces of different weight or age which all fit the given skull equally well.

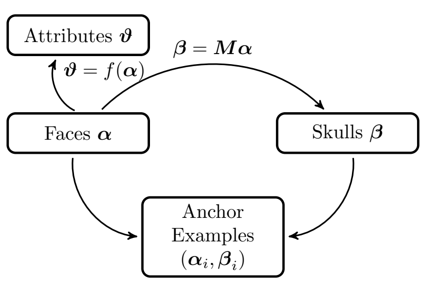
The problem is formulated as a minimization problem for the coefficients $\boldsymbol{\alpha}$ of the face model. The coefficients which best fit a set of skull coefficients $\boldsymbol{\beta}$ and attributes $\boldsymbol{\vartheta}$ are sought as the minimum of a compound functional:

$$E(\boldsymbol{\alpha}) = E_s(\boldsymbol{\alpha}, \boldsymbol{\beta}) + \lambda_1 E_a(\boldsymbol{\alpha}, \boldsymbol{\vartheta}) + \lambda_2 E_p(\boldsymbol{\alpha}). \tag{4}$$

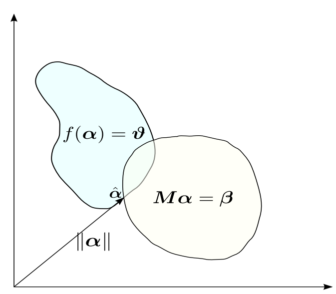
λ_1 and λ_2 are weights to balance the influence of the three terms of the functional:

- The *skull error* $E_s(\boldsymbol{\alpha}, \boldsymbol{\beta})$ describes how well the predicted face fits the given skull model coefficients $\boldsymbol{\beta}$.
- The *attribute error* $E_a(\boldsymbol{\alpha}, \boldsymbol{\vartheta})$ measures how well the predicted face coefficients $\boldsymbol{\alpha}$ match the user defined attributes $\boldsymbol{\vartheta}$.
- The *prior* $E_p(\boldsymbol{\alpha})$ quantifies the probability that the predicted $\boldsymbol{\alpha}$ represents a valid face. It has a regularizing effect and reduces overfitting.

The goal is to find coefficients $\boldsymbol{\alpha}$ that minimize all three terms simultaneously, as illustrated in Figure 1b. In the following subsections we will discuss these three terms in more depth.



(a) Schema of the data used



(b) Face model parameters

Fig. 1: (a) Face and skull models are described by parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, face attributes by $\boldsymbol{\vartheta}$. For the anchor examples, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are known. The mappings $a = f(\boldsymbol{\alpha})$ and $\boldsymbol{\beta} = \boldsymbol{M}\boldsymbol{\alpha}$ are learned from the data. (b) It is assumed that several faces $\boldsymbol{\alpha}$ fit a given skull $\boldsymbol{\beta}$ as well as the attributes $\boldsymbol{\vartheta}$. We search for the $\hat{\boldsymbol{\alpha}}$ with minimal norm conforming to both requirements

3.1 Linear Skull Predictor

As the most important step of our method, we wish to establish a relationship between the previously independent skull and face model. This is achieved by learning the relation from the face and skulls surfaces given in the training examples. We can fit both models to these "anchor examples" to get n pairs of corresponding parameters $\{(\boldsymbol{\alpha}_i, \boldsymbol{\beta}_i) \,|\, i = 1, \ldots, n\}$, cf. Section 2.3. For each individual i, $\boldsymbol{\alpha}_i \in \mathbb{R}^{m_f}$ are the parameters of the face model, and $\boldsymbol{\beta}_i \in \mathbb{R}^{m_s}$ those of the corresponding skull in the skull model. Using these pairs as training data, we wish to learn a mapping $\boldsymbol{M}$, from the face parameters to the skull parameters, i.e. $\boldsymbol{M\alpha} = \boldsymbol{\beta}$. While in principle this can be achieved with any machine learning approach, we learn a linear mapping. Preferring linear over more complicated mappings has two reasons. First, assuming that an observed face surface can be well represented as a linear combination of training examples, we would expect the underlying skull to be the same combination of the skulls of the training examples, which leads to a linear mapping. Secondly, due to the limited number of training examples, we wish to use a relatively simple model.

We now expand the above argument that if a face is well represented as a combination of example faces then its skull should be well represented by the same combination of the corresponding example skulls. For the anchor examples, for which we have both face and skull data, we write the face model parameters $\boldsymbol{\alpha}_i$ as a Matrix $\boldsymbol{A} := [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n] \in \mathbb{R}^{m_f \times n}$ and the skull model parameters $\boldsymbol{\beta}_i$ as $\boldsymbol{B} := [\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n] \in \mathbb{R}^{m_s \times n}$. To predict skull parameters $\hat{\boldsymbol{\beta}}$ from face parameters $\boldsymbol{\alpha}$ of a newly observed face we first find a linear combination $\hat{\boldsymbol{\alpha}} = \boldsymbol{Ac}$ of example face parameters best approximating $\boldsymbol{\alpha}$. This is done by projecting $\boldsymbol{\alpha}$ into the space of the example faces:

$$\boldsymbol{c} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\alpha} = \arg\min_{\boldsymbol{c}} \|\boldsymbol{Ac} - \boldsymbol{\alpha}\|^2. \tag{5}$$

the coefficient $\boldsymbol{c}$ are then used to generate the corresponding skull parameters

$$\hat{\boldsymbol{\beta}} = \boldsymbol{Bc} = \boldsymbol{B}\,(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{\alpha} =: \boldsymbol{M\alpha}. \tag{6}$$

As we have relatively few examples, it is necessary to introduce some regularisation in the projection. Therefore we change the above to:

$$\hat{\boldsymbol{\beta}} = \boldsymbol{Bc} = \boldsymbol{B}\,(\boldsymbol{A}^T\boldsymbol{A} + \lambda\boldsymbol{I})^{-1}\boldsymbol{A}^T\boldsymbol{\alpha} =: \boldsymbol{M\alpha}. \tag{7}$$

The mapping matrix $\boldsymbol{M} = \boldsymbol{B}\,(\boldsymbol{A}^T\boldsymbol{A} + \lambda\boldsymbol{I})^{-1}\boldsymbol{A}^T$ can equivalently be determined by ridge regression from face parameters to skull parameters:

$$\boldsymbol{M} = \arg\min_{\boldsymbol{M}} \; \|\boldsymbol{MA} - \boldsymbol{B}\|_F^2 + \lambda^2\|\boldsymbol{M}\|_F^2, \tag{8}$$

where $\|\cdot\|_F$ is the Frobenius norm. For more details on ridge regression, see e.g. [10]. The mapping $\boldsymbol{M}$ is calculated only once from the training data and can then be used for all subsequent reconstructions. By exchanging $\boldsymbol{A}$ and $\boldsymbol{B}$, we can exchange the role of faces and skulls and make a prediction in the opposite direction. For our overall error function Equation (4) however, we need to evaluate how well the estimated face coefficients $\boldsymbol{\alpha}$ fit the given skull coefficients $\boldsymbol{\beta}$ in skull space and therefore calculate the mapping $\boldsymbol{M}$ from face to skull coefficients. We define the error term $E_s(\boldsymbol{\alpha})$ in Equation (4) as:

$$E_s(\boldsymbol{\alpha}) := \|\boldsymbol{M}\boldsymbol{\alpha} - \boldsymbol{\beta}\|^2. \tag{9}$$

It measures how well the face coefficients $\boldsymbol{\alpha}$, or rather their mapping $\hat{\boldsymbol{\beta}} = \boldsymbol{M}\boldsymbol{\alpha}$, fit the input skull coefficients $\boldsymbol{\beta}$. This is the Mahalanobis distance in skull space, which is commonly used as a measurement for the similarity of two shapes.

3.2 Attribute Prediction from Face Coefficients

The attribute error term $E_a(\boldsymbol{\alpha}, \boldsymbol{\vartheta})$ measures how well the set of face parameters $\boldsymbol{\alpha}$ matches the chosen attributes $\boldsymbol{\vartheta}$. We relate these different values to each other by learning a function f mapping the face coefficients $\boldsymbol{\alpha}$ to the corresponding attributes $\boldsymbol{\vartheta} = f(\boldsymbol{\alpha})$. Similar to the skull prediction, we use a training set with known matching parameter pairs $(\boldsymbol{\alpha}_i, \boldsymbol{\vartheta}_i)$ to train the function f. As the attributes are known for all 840 face examples used to build the face model, we have a much larger training set and can also use nonlinear functions to learn this relationship. Notably, we train a support vector regression with radial basis function (RBF) kernels. We use the LIBSVM implementation for ν-Support Vector Regression [11] to find the parameters α_j, $\alpha_j^*, b \in \mathbb{R}$ of the RBF support vector regression function:

$$f(\boldsymbol{x}) = \sum_{j=1}^{l} (-\alpha_j + \alpha_j^*) e^{-\gamma \|\boldsymbol{x}_j - \boldsymbol{x}\|^2} + b. \tag{10}$$

Here, l is the number of face examples $\boldsymbol{x_j}$. The kernel width γ and the upper bound for α_i and α_i^* are determined by grid search and ten fold cross validation. For each recorded attribute, a regression function f_i is learned. The attribute error function $E_a(\boldsymbol{\alpha}, \boldsymbol{\vartheta})$ is then defined as:

$$E_a(\boldsymbol{\alpha}, \boldsymbol{\vartheta}) := \sum_{i \in I} (w_i (f_i(\boldsymbol{\alpha}) - \vartheta_i))^2, \tag{11}$$

where I is an index set for the different attributes and the w_i are normalization factors for the value ranges of the different attributes.

3.3 Minimization and Face Prediction

We are interested only in solutions $\boldsymbol{\alpha}$, which represent a valid face. The last term $E_p(\boldsymbol{\alpha}) := \|\boldsymbol{\alpha}\|^2$ therefore penalizes unlikely faces (cf. Equation (2)). To find the minimum of the full functional (4) we use a conjugated gradient optimization method with

$$\nabla E(\boldsymbol{\alpha}) = 2(\boldsymbol{M}^T\boldsymbol{M}\boldsymbol{\alpha} - \boldsymbol{M}^T\boldsymbol{\beta}) + 2\lambda_1 \sum_i w_i^2 (f_i(\boldsymbol{\alpha}) - \vartheta_i) f_i'(\boldsymbol{\alpha}) + 2\lambda_2 \boldsymbol{\alpha}, \tag{12}$$

where $f_i'(\boldsymbol{x})$ is the derivative of the SVR function Equation (10). Note that the term E_a in (4) is non-linear, and hence it is important to choose a good initial solution. Such an initial solution can be obtained by direct prediction of the coefficients $\boldsymbol{\alpha}$ from $\boldsymbol{\beta}$, in the same manner as described in section 3.1.

4 Results

For our experiments, we have used the data sets introduced in Section 2.2. The 840 face scans were brought into correspondence with a non-rigid iterative closest point algorithm [12] and the 20 skull surfaces were brought into correspondence with a variational optical flow approach [13]. In the following we present the experimental results for the different parts of our algorithm individually. Finally we show results where we manipulate the attributes for the obtained reconstruction.

4.1 Skull and Face Prediction without Attributes

First we evaluate the ability of the linear skull predictor introduced in Section 3.1 to reconstruct a skull from given face parameters. We conducted a leave-one-out experiment, comparing the prediction $\boldsymbol{M}$ trained on all but one of the anchor examples to the ground truth given by this left out example. In this experiment a parameter selection is used to determine a good regularization parameter λ. The best and the worst results are displayed in Figure 2a. For the prediction error of the skulls we obtained the mean absolute error (MAE) 1.2416 mm and its standard deviation (STD) of 1.1768 mm.

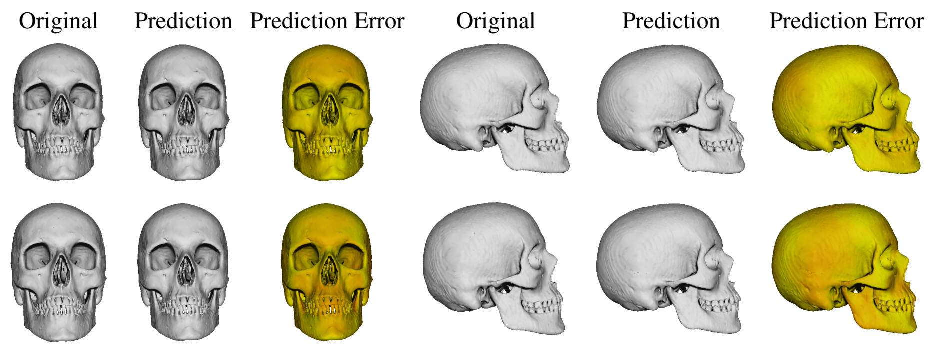
Further, we tested the face prediction presented in Section 3.4, but still without attribute manipulation, i.e. we set the weighting parameter for the attribute term λ_1 to zero. The best and worst result are shown in Figure 2b. We observe that the largest reconstruction errors occur in places where the soft tissue thickness can vary, whereas the eye and mouth area are well reconstructed even in the bad examples. Errors in the forehead and neck are mostly due to the model's boundary conditions. While it is easy to recognize the best predicted face, the worst reconstruction is not close enough to the ground truth to be able to recognize the person's face anymore. We obtained 2.8499 mm MAE and 2.4214 mm STD.

4.2 Attribute Prediction

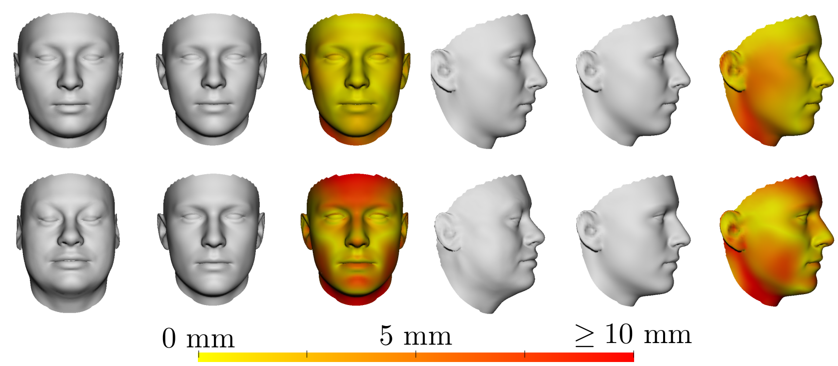
Before performing the full face prediction with attribute manipulation, we tested the performance of the attribute prediction function introduced in Section 3.2. Figure 3 shows the true values plotted against the predicted values. A perfect prediction would produce values only on the diagonal. The values for weight, height and age are sufficiently close to the diagonal, while the values for sex show a good approximation of the binary attribute male/female with a continuous variable. We obtained the MAE $[0.24, 3.64, 3.21, 3.31]$ with STD $[0.25, 4.49, 4.12, 4.18]$ for sex, weight, height and age.

4.3 Face Prediction

To evaluate the face prediction results we estimated for each examples of our MRI data set the corresponding faces (figure 2b). To separate the training from the test set we used again a leave one out scheme. To obtain an optimal reconstruction, we estimated the attributes for the face coefficients using the trained regression function. For each of the examples we predicted different results with varying attributes. Examples are shown in figure 4 and 5, where we show results for the most interesting attributes, weight and age.

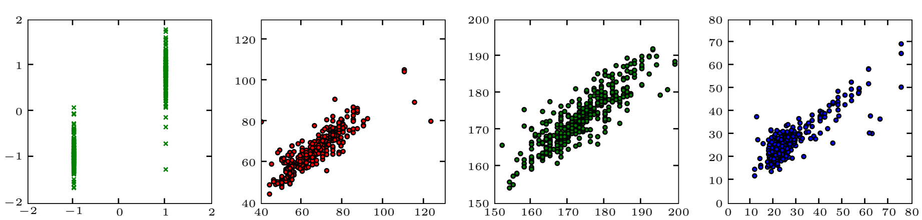


(a) Skull prediction: Best and worst example



(b) Face prediction (without attribute manipulation): Best and worst example

Fig. 2: Results of skulls predicted from faces and vice versa. In both cases, the best and worst results in terms of the Mahalanobis norm error were selected. The color-coded prediction error is the per-vertex L^2-error orthogonal to the surface. For the face prediction large errors occur at the cheeks where the soft tissue thickness depends strongly on the body weight and age.



(a) Sex (b) Weight (c) Height (d) Age

Fig. 3: Support Vector Regression results obtained by 10 fold cross validation on the face database. Predicted (y-axis) sex (1,-1 for male and female), weight, height and age plotted against the true value (x-axis).

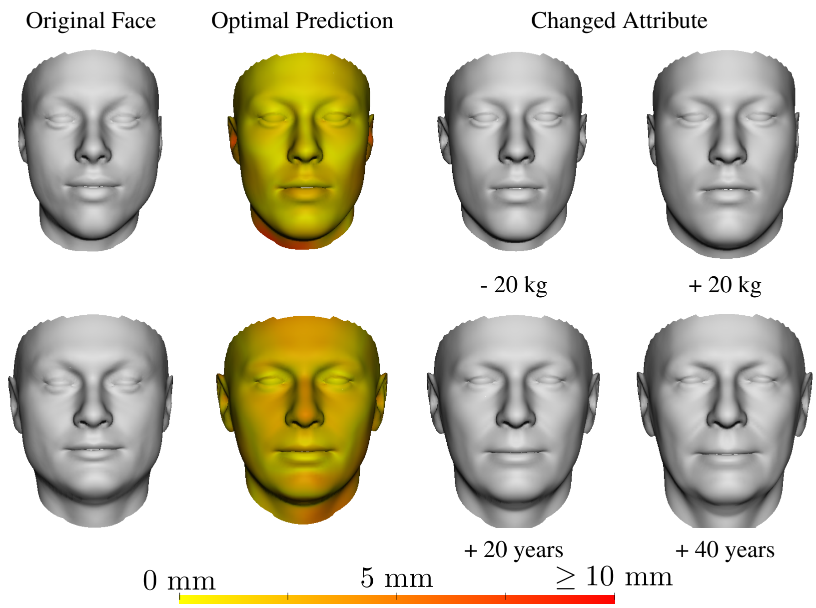


Fig. 4: Results of the face prediction with attribute manipulation of the original faces (first column). The second column shows the reconstruction with the optimally estimated attributes. The renderings in the right column are obtained by varying the attributes weight and age.

5 Conclusion

While a considerable amount of research has been devoted to face reconstruction, it is still arguable whether either of the techniques produces reliable results. Indeed, in a study performed in 2001, Stephan et al. [14] conclude that among 4 standard techniques for facial reconstruction, the 3D American method was the only method that gave identification rates slightly above chance rate. Our results confirm, that even though the prediction are close in terms of the average error, the individual is difficult to recognize. By constraining the result to satisfy certain attributes, the reconstruction comes perceptually closer to the original face.

While the experimental results show the feasibility of our method, we see the biggest advantage of our method in the formulation of the problem in terms of finding a relationship among separate shape model parameters. This formulation does allow us to use prior knowledge about faces and skulls that can be acquired independently, using the suitable acquisition method for each model. Furthermore, the learning approach allows the use of the wide variety of algorithms developed in the field to find statistical dependencies among the model coefficients. While, due to the limited number of training examples, we used a simple linear regression function, we believe that the results can be improved using more data and more sophisticated methods such as canonical correlation analysis, to single out parameters which strongly correlate between the ex-

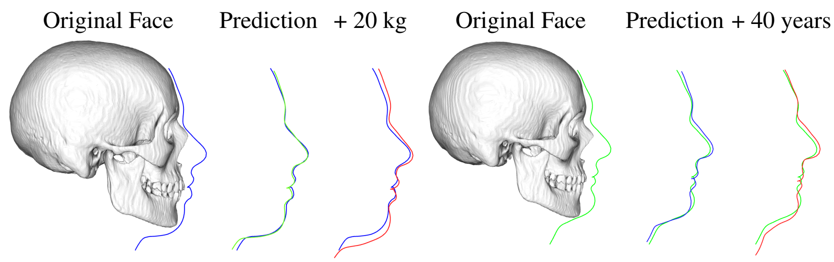


Fig. 5: Horizontal cuts to visualize the prediction results.

amples for predicting the shapes. The other parameters could then be set depending on the specified attributes. Investigating this possibility will be the subject of future work.

References

1. Wilkinson, C.: Computerized forensic facial reconstruction. Forensic Science, Medicine, and Pathology **1**(3) (2005) 173–177
2. Starbuck, J., Ward, R.: The affect of tissue depth variation on craniofacial reconstructions. Forensic Science International **172**(2-3) (2007) 130–136
3. Muller, J., Mang, A., Buzug, T.: A template-deformation method for facial reproduction. In: Image and Signal Processing and Analysis, 2005. ISPA 2005. Proceedings of the 4th International Symposium on. (2005) 359–364
4. Kähler, K., Haber, J., Seidel, H.: Reanimating the dead: reconstruction of expressive faces from skull data. ACM Transactions on Graphics (TOG) **22**(3) (2003) 554–561
5. Pei, Y., Zha, H., Yuan, Z.: The Craniofacial Reconstruction from the Local Structural Diversity of Skulls. In: Computer Graphics Forum. Volume 27., Blackwell Publishing Ltd (2008)
6. Claes, P., Vandermeulen, D., De Greef, S., Willems, G., Suetens, P.: Craniofacial reconstruction using a combined statistical model of face shape and soft tissue depths: Methodology and validation. Forensic Science International **159** (2006) 147–158
7. Berar, M., Desvignes, M., Bailly, G., Payan, Y.: 3D statistical facial reconstruction. In: Image and Signal Processing and Analysis, ISPA 2005. (2005) 365–370
8. Tu, P., Hartley, R., Lorensen, W., Allyassin, M., Gupta, R., Heier, L.: Face reconstructions using flesh deformation modes. International Association for Craniofacial Identification (2000)
9. anonymised: A hierarchical, multi-resolution approach for model-based skull-segmentation in mri volumes. Technical report, anonymous (2009)
10. Shawe-Taylor, J., Cristianini, N.: Kernel methods for pattern analysis. Cambridge University Press (2004)
11. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. (2001) Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.
12. Amberg, B., Romdhani, S., Vetter, T.: Optimal step nonrigid ICP algorithms for surface registration. IEEE CVPR (Jun 2007) 1–8
13. Dedner, A., Lüthi, M., Albrecht, T., Vetter, T.: Curvature guided level set registration using adaptive finite elements. In: Pattern Recognition. (2007) 527–536
14. Stephan, C., Henneberg, M.: Building faces from dry skulls: are they recognized above chance rates? Journal of Forensic Sciences **46**(3) (2001) 432–440