Face Identification across Different Poses and Illuminations with a 3D Morphable Model

V. Blanz, S. Romdhani, and T. Vetter
University of Freiburg
Georges-Köhler-Allee 52, 79110 Freiburg, Germany
{volker, romdhani, vetter}@informatik.uni-freiburg.de

Abstract

We present a novel approach for recognizing faces in images taken from different directions and under different illumination. The method is based on a 3D morphable face model that encodes shape and texture in terms of model parameters, and an algorithm that recovers these parameters from a single image of a face. For face identification, we use the shape and texture parameters of the model that are separated from imaging parameters, such as pose and illumination. In addition to the identity, the system provides a measure of confidence. We report experimental results for more than 4000 images from the publicly available CMU-PIE database.

1 Introduction

A number of algorithms have been developed for face recognition from fixed viewpoints, yet few attempts have been made to tackle the problem of combined variation of pose and illumination (an overview is given in [15].) To handle the extreme image variations induced by these parameters, one common approach taken by various groups is to use generative image models [6, 10, 14, 5]. For image analysis, the general strategy of all these techniques is to fit the generative model to a novel image, thereby parameterizing it in terms of the model.

In order to make identification independent of imaging conditions, the goal is to separate intrinsic model parameters of the face from extrinsic imaging parameters. Many view-based approaches use statistical techniques to address this problem. Head poses that range from frontal to profile views need to be covered in view-based methods by a set of separate models for different views [3]. In another approach, given several front views with different illumination directions of each person, a set of models for new poses can be generated synthetically and then combined to cover limited rotations in azimuth of up to $24°$ [5].

In our approach, the separation of intrinsic and extrinsic parameters is achieved by taking the model-based approach to its extreme in an explicit simulation of the process of image formation using 3D computer graphics technology. We apply a three-dimensional morphable face model that has previously been introduced for computer graphics, and an algorithm to fit this model to images [2]. From a single image, the algorithm estimates facial shape and texture, along with pose, illumination, and camera parameters such as color contrast. Our 3D model covers all head poses and a wide range of illumination conditions, and it considers specular reflection and cast shadows. For fitting the model to an image, the system currently requires approximate prior information about these external conditions.

In the following section, we summarize the concept of the morphable face model. In Section 3, we describe an algorithm for recovering model parameters from images. Finally, we present results obtained with the CMU-PIE database of face images [12].

2 Morphable Model of 3D Faces

Generalizing the well-known morphing between pairs of three-dimensional objects, the morphable face model is based on a vector space representation of faces [14]. In this vector space, any convex combination of shape and texture vectors of a set of examples describes a realistic human face. The following paragraphs describe an automated technique to derive a morphable model from a set of laser scans, and give a definition of shape and texture vectors.

2.1 Database of 3D Laser Scans

The database of laser scans used in this study contains scans of 100 males and 100 females recorded with a $Cyberware^{TM}$ 3030PS scanner. Scans are stored in cylindrical coordinates relative to a vertical axis. In angular steps ϕ and vertical steps h, at a spacing of $0.7°$ and $0.615\mathrm{mm}$, the device measures radius r, along with red, green and blue components of surface texture R, G, B:

$$\mathbf{I}(h,\phi) = (r(h,\phi), R(h,\phi), G(h,\phi), B(h,\phi))^T \qquad (1)$$

All heads were consistently aligned in 3D space with the method of 3D3D absolute orientation [7]. Since surface data are unavailable for hair, the back of the head was removed from each scan, using an interactive tool.

2.2 Correspondence based on Optic Flow

The core step of building a morphable face model is to establish dense point-to-point correspondence between each face and a reference face, which can be a scan from the database or any other 3D face model. Dense correspondence is given by a vector field $\mathbf{v}(h,\phi) = (\Delta h(h,\phi),\, \Delta\phi(h,\phi))^T$ such that each point $\mathbf{I}_1(h,\phi)$ in the first scan corresponds to the point $\mathbf{I}_2(h+\Delta h,\, \phi+\Delta\phi)$ in the second scan. To find this vector field, we extended an optic flow algorithm [1] from grey-level images $I(x,y)$ to vector-valued arrays $\mathbf{I}(h,\phi)$, replacing products of grey values $I_1(x,y)\cdot I_2(x,y)$ in the algorithm by scalar products

$$\langle \mathbf{I}_1, \mathbf{I}_2 \rangle = w_r r_1 r_2 + w_R R_1 R_2 + w_G G_1 G_2 + w_B B_1 B_2 \quad (2)$$

with weight factors w_r, w_R, w_G, w_B that compensate for different variations within the radius and texture data. The coordinates and texture values of all n vertices of the reference face (in our model, $n = 75972$) are concatenated to shape and texture vectors

$$\mathbf{S}_0 = (x_1, y_1, z_1, x_2, \ldots, x_n, y_n, z_n)^T, \quad (3)$$

$$\mathbf{T}_0 = (R_1, G_1, B_1, R_2, \ldots, R_n, G_n, B_n)^T. \quad (4)$$

Vectors $\mathbf{S}_i$ and $\mathbf{T}_i$ of the examples $i = 1 \ldots m$ in the database are formed in a consistent way using the flow field $\mathbf{v}(h,\phi)$ from the reference face to face i. Convex combinations of the examples produce novel shape and texture vectors $\mathbf{S}$ and $\mathbf{T}$. Previous results [2] indicate that shape and texture can be combined independently:

$$\mathbf{S} = \sum_{i=1}^{m} a_i \mathbf{S}_i, \quad \mathbf{T} = \sum_{i=1}^{m} b_i \mathbf{T}_i. \quad (5)$$

To avoid changes in overall size and brightness, a_i and b_i should sum to 1. The additional constraints $a_i, b_i \in [0,1]$ imposed on convex combinations will be replaced by a probabilistic criterion in the next section.

2.3 Principal Component Analysis

We perform a Principal Component Analysis (PCA, see [8]) separately on the shape and texture vectors $\mathbf{S}_i$ and $\mathbf{T}_i$, ignoring the correlation between shape and texture data. For shape, subtracting the average $\overline{\mathbf{s}}$,

$$\mathbf{a}_i = \mathbf{S}_i - \overline{\mathbf{s}}, \quad \text{with} \quad \overline{\mathbf{s}} = \frac{1}{\mathrm{m}} \sum_{\mathrm{i}=1}^{\mathrm{m}} \mathbf{S}_\mathrm{i}$$

a data matrix $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$ can be defined. We calculate the eigenvectors of the covariance matrix $\mathbf{C} = \frac{1}{m}\mathbf{A}\mathbf{A}^T$ by a Singular Value Decomposition of $\mathbf{A}$. The eigenvalues of $\mathbf{C}$, $\sigma_{S,1}^2 \geq \sigma_{S,2}^2 \geq \ldots$, describe the variance within the data along each eigenvector $\mathbf{s}_1, \mathbf{s}_2, \ldots$. By the same procedure, we obtain texture variances $\sigma_{T,i}^2$ and eigenvectors $\mathbf{t}_i$. The eigenvectors form an orthogonal basis,

$$\mathbf{S} = \overline{\mathbf{s}} + \sum_{i=1}^{m-1} \alpha_i \cdot \mathbf{s}_i, \qquad \mathbf{T} = \overline{\mathbf{t}} + \sum_{i=1}^{m-1} \beta_i \cdot \mathbf{t}_i \quad (6)$$

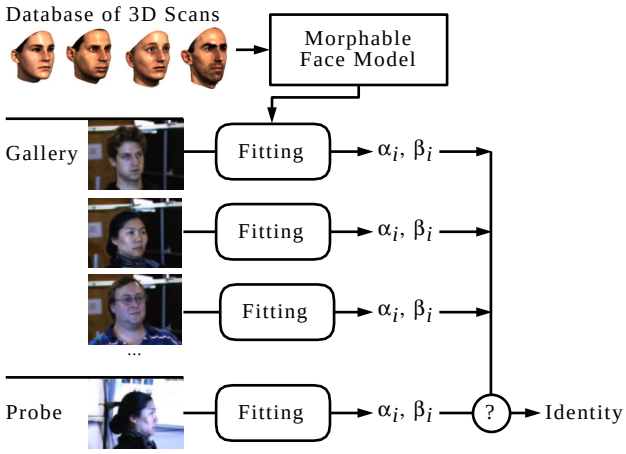

Figure 1. Derived from a database of laser scans, the three-dimensional morphable face model is used to encode gallery and probe images. For identification, the model coefficients of the probe image are compared with the coefficients of all gallery images.

Moreover, PCA gives an estimate of the probability density function within face space, based on the set of examples

$$p_S(\mathbf{S}) \sim e^{-\frac{1}{2}\sum_i \frac{\alpha_i^2}{\sigma_{S,i}^2}}, \qquad p_T(\mathbf{T}) \sim e^{-\frac{1}{2}\sum_i \frac{\beta_i^2}{\sigma_{T,i}^2}}. \quad (7)$$

3 Model-Based Image Analysis

In an analysis-by-synthesis loop, the morphable face model can be fitted to a novel face shown in an input image $\mathbf{I}_{input}(x,y)$. The goal of the image analysis is to find model parameters α_i and β_i, and face position, orientation and illumination such that the model, rendered by computer graphics algorithms, produces an image as close as possible to the input image (Figure 2, illustrated with linear combinations according to (5) rather than (6) for visualization.)

We first summarize the algorithms and parameters involved in creating an image from the model, and then discuss how the model is fitted to an input image. Finally, we describe our identification criterion.

3.1 Image Synthesis

The three-dimensional positions and the color of the model's vertices are defined by the coefficients α_i and β_i in (6). Rendering an image includes the following steps:

Image positions of vertices: A rigid transformation maps the object-centered coordinates $\mathbf{x}_k = (x_k, y_k, z_k)^T$ of each vertex k to a position relative to the camera:

$$(w_{x,k}\,,\, w_{y,k}\,,\, w_{z,k})^T = s \cdot \mathbf{R}_\gamma \mathbf{R}_\theta \mathbf{R}_\phi \mathbf{x}_k + \mathbf{t}_w. \quad (8)$$

The angles ϕ and θ control in-depth rotations around the vertical and horizontal axis, and γ defines a rotation around the camera axis. s is a scaling factor, and $\mathbf{t}_w$ a spatial shift.

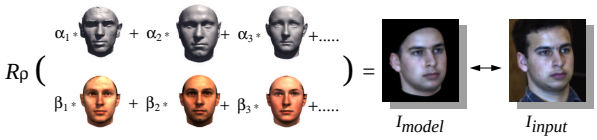

Figure 2. The goal of the fitting process is to find shape and texture coefficients α and β such that rendering $\mathbf{R}_\rho$ produces an image $\mathbf{I}_{\mathbf{model}}$ that is as similar as possible to $\mathbf{I}_{\mathbf{input}}$.

In a perspective projection, the coordinates of vertex k are mapped to image plane coordinates $p_{x,k}$, $p_{y,k}$:

$$p_{x,k} = P_x + f\frac{w_{x,k}}{w_{z,k}}\,, \qquad p_{y,k} = P_y - f\frac{w_{y,k}}{w_{z,k}}\,. \tag{9}$$

f is the focal length of the camera which is located in the origin, and $(P_x\,, P_y)$ defines the image-plane position of the optical axis (principal point).

Illumination and Color: Shading of surfaces depends on the direction of the surface normals $\mathbf{n}$. In the fitting process, we only consider the centers of triangles, most of which are about 0.2mm^2 in size. Normals at triangles' centers are computed by a vector product of the edges, normalized to unit length, and rotated along with the head (Equation 8).

The illumination model of Phong (see [4]) approximately describes the diffuse and specular reflection on a surface. On each vertex k, the red channel is

$$L_{r,k} = R_k \cdot L_{r,amb} + R_k \cdot L_{r,dir} \cdot \langle \mathbf{n}_k, \mathbf{l} \rangle + k_s \cdot L_{r,dir} \langle \mathbf{r}_k, \hat{\mathbf{v}}_k \rangle^\nu$$

where R_k is the red component of the diffuse reflection coefficient stored in the texture vector $\mathbf{T}$, $L_{r,amb}$ and $L_{r,dir}$ are the red intensities of the ambient and directed light, $\mathbf{l}$ is the direction of illumination, k_s is the specular reflectance, ν defines the angular distribution of the specularities, $\hat{\mathbf{v}}_k$ is the viewing direction, and $\mathbf{r}_k = 2 \cdot \langle \mathbf{n}_k, \mathbf{l} \rangle \, \mathbf{n}_k - \mathbf{l}$ is the direction of maximum specular reflection [4].

Input images may vary a lot with respect to the overall tone of color. In order to be able to handle a variety of color images, we transform color values with respect to color contrast, as well as gain and offset in each channel. The transformed colors I_r, I_g and I_b are drawn at a position (p_x, p_y) in the final image $\mathbf{I}_{model}$.

3.2 Fitting the Model to an Image

In the process of fitting the model to a novel image, not only the shape and texture coefficients α_i and β_i are optimized, but also the following rendering parameters, which we concatenate into a vector $\boldsymbol{\rho}$: The angles ϕ, θ and γ, the head position in the image plane, controlled by (P_x, P_y), size s, color and intensity of the light sources $L_{r,amb}$, $L_{g,amb}$, $L_{b,amb}$ and $L_{r,dir}$, $L_{g,dir}$, $L_{b,dir}$, as well as color contrast, and gain and offset of colors.

Other parameters, such as the focal length f of the camera, the 3D translation $\mathbf{t}_w$, the direction $\mathbf{l}$ of the light, and the specular reflectance parameters are kept fixed.

The optimization algorithm starts from the average face at a position and orientation roughly aligned with the face in the image. The prealignment is currently done manually.

The primary goal in analyzing a face is to minimize the sum of square differences over all color channels and all pixels in the input image and the synthetic reconstruction,

$$E_I = \sum_{x,y} \left\| \mathbf{I}_{input}(x,y) - \mathbf{I}_{model}(x,y) \right\|^2 . \tag{10}$$

For Gaussian pixel noise with a standard deviation σ_N, the likelihood of observing $\mathbf{I}_{input}$, given α, β, ρ, is $P(\mathbf{I}_{input} | \alpha, \beta, \rho) \sim exp(\frac{-1}{2\sigma_N^2} \cdot E_I)$.

In order to achieve plausible results, we use a maximum a posteriori estimator (MAP), minimizing $E = -ln(P(\mathbf{I}_{input} | \alpha, \beta, \rho) \cdot p(\alpha, \beta, \rho))$ [2] . Assuming that α and β are independent, the prior probability $p(\alpha, \beta, \rho)$ is the product of the probabilities in Equation (7) and a normal distribution for ρ_i, using the starting values for $\overline{\rho}_i$ and ad hoc values for $\sigma_{R,i}$. The overall cost function to be minimized is then

$$E = \frac{1}{\sigma_N^2} E_I + \sum_i \frac{\alpha_i^2}{\sigma_{S,i}^2} + \sum_i \frac{\beta_i^2}{\sigma_{T,i}^2} + \sum_i \frac{(\rho_i - \overline{\rho}_i)^2}{\sigma_{R,i}^2}. \tag{11}$$

For each iteration of the optimization process, the fitting algorithm analytically computes the gradient of the cost function and then updates the parameters:

$$\alpha_i \mapsto \alpha_i - \lambda \frac{\partial E}{\alpha_i}, \; \beta_i \mapsto \beta_i - \lambda \frac{\partial E}{\beta_i}, \; \rho_i \mapsto \rho_i - \lambda \frac{\partial E}{\rho_i}. \tag{12}$$

Since contributions of the pixels of the entire image might be redundant, we use a modification of a stochastic gradient descent algorithm [9]. At each iteration, the fitting algorithm selects a subset of 40 random triangles, and evaluates E_I and its gradient only at their centers. To make sure that the expectation value of this reduced cost function equals the entire function E_I on the face, we select triangles with probabilities that are proportional to the areas they cover in the image. Areas are computed at the beginning of the optimization, and once every few thousand iterations. At the same time, a z-buffer method discards all triangles that are occluded in the current estimate of shape and orientation.

The first iterations only optimize the first parameters $\alpha_i, \beta_i, i \in \{1, ..., 10\}$ and all parameters ρ_i, at a high value of σ_N. Later in the optimization, the algorithm considers more and more coefficients, and σ_N is decreased. While PCA of a database of 200 faces provides 199 coefficients α_i, β_i, we restrict the process to the most relevant 99.

In order to recover details of the face more precisely, we defined the regions of the eyes, nose, mouth, and the surrounding face segment on the reference face [2]. Thus, each vertex of the morphable model is assigned to one of the four segments $s1$ ... $s4$. After the entire face model has been globally fitted to the image, these regions are optimized separately, using independent linear combinations (6) and evaluating E_I only on that segment. Currently, the fitting process takes 40 minutes on a Pentium III, 800MHz.

3.3 Identification Criterion

As a result of the fitting procedure, we obtain a set of parameters $\boldsymbol{\alpha}_g = (\alpha_1, \ldots, \alpha_{99})^T$ and $\boldsymbol{\beta}_g = (\beta_1, \ldots, \beta_{99})^T$ for the global result, and one set $\boldsymbol{\alpha}_{s1}$, $\boldsymbol{\beta}_{s1}$ ... $\boldsymbol{\alpha}_{s4}$, $\boldsymbol{\beta}_{s4}$ for each of the face segments. All of these parameters are combined for a comparison of faces.

For each set of parameters, we define a scalar product

$$\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}' \rangle_M = \sum_i \frac{\alpha_i \cdot \alpha_i'}{\sigma_{S,i}^2}, \quad \langle \boldsymbol{\beta}, \boldsymbol{\beta}' \rangle_M = \sum_i \frac{\beta_i \cdot \beta_i'}{\sigma_{T,i}^2}.$$

Hence, the Mahalanobis distance [8] from the average is $\|\boldsymbol{\alpha}\|_M^2 = \langle \boldsymbol{\alpha}, \boldsymbol{\alpha} \rangle_M$ and $\|\boldsymbol{\beta}\|_M^2 = \langle \boldsymbol{\beta}, \boldsymbol{\beta} \rangle_M$.

The similarity d between two faces can be measured in terms of Mahalanobis distances. However, it proved to be more reliable (cf. [11]) to use

$$d = \sum_{g,s_1,s_2,s_3,s_4} \left(\frac{\langle \boldsymbol{\alpha}, \boldsymbol{\alpha}' \rangle_M}{\|\boldsymbol{\alpha}\|_M \cdot \|\boldsymbol{\alpha}'\|_M} + \frac{\langle \boldsymbol{\beta}, \boldsymbol{\beta}' \rangle_M}{\|\boldsymbol{\beta}\|_M \cdot \|\boldsymbol{\beta}'\|_M} \right). \quad (13)$$

4 Experiments on Model Fitting

In this and the next sections, the fitting and identification performances are investigated on the PIE database from CMU [12]. We selected the portion of the database which presents variation of both pose and directed illumination and with an ambient light. This portion includes images of 68 individuals at 3 poses and illuminated from 22 different directions (66 images per individual). The individuals in the PIE database are not contained in our set of 3D scans.

As explained in Section 3.2, initial values of the rendering parameters must be provided to the image analysis algorithm. Some of these are computed automatically from the geometrical information provided by the CMU database: The 3D position of the cameras, the light sources and the head were recorded at the time of the data acquisition [12]. These data allow us to recover an initial estimate of the pose angles ϕ and the direction of the light $\mathbf{l}$. On the other hand, translation and size parameters, (P_x, P_y) and s, must be manually adjusted (once per view of each person). The other parameters, θ, γ, f, $L_{r,amb}$, $L_{g,amb}$,$L_{b,amb}$, $L_{r,dir}$, $L_{g,dir}$, $L_{b,dir}$, color contrast, gains and offsets, were given generic values for the entire set of images. Except for the focal length f and for the light direction $\mathbf{l}$, the parameters are refined by the image analysis algorithm.

Figure 3 shows examples of initial parameters and of parameters recovered by the algorithm. As the fifth column of this figure shows, the iterative fitting does not always converge to an acceptable solution. After fitting the model to all 4488 images on several PCs, results were assessed visually by an operator. Percentages of acceptable fitting results are shown in Table 1.

Our experiments indicate that it is easier to fit a front face than a profile face. Glasses affect the quality of results considerably (28 out of 68 individuals wear glasses), in particular on the side view where glasses take more space than from the front view. Ignoring those points with highest contribution to E_I in the fitting procedure [9] makes our algorithm robust with respect to glasses and beards. However, this method may reduce quality for difficult poses and illuminations, and is therefore not applied here.

	front	side	profile
mean w/o glasses	94 %	91 %	80 %
mean with glasses	86 %	65 %	72 %
mean	91 %	80 %	77 %
worst and best illum.	68-98%	41-95 %	14-89 %

Table 1. Percentage of successful fitting for 3 poses averaged over all lighting conditions and individuals.

	+	-
Vis. assessed +	95 %	5 %
Vis. assessed -	13 %	87 %

	+	-
+	93 %	7 %
-	21 %	79 %

Table 2. Results for training (left) and validation (right) of the Fitting Score SVM.

The quality of results varies across different illumination conditions (Table 1, last row). It is consistently low for some extreme conditions where most of the face is dark except for strong specularities on the forehead and the chin (see third column of Figure 3).

4.1 Fitting Score

In this section, we present an automated technique for assessing the quality of fitting in terms of a *Fitting Score* (FS). We demonstrate in the next section that the FS is correlated with identification performance and may be used as a confidence measure.

A fitting score can be derived from the image error and from the model coefficients of each fitted segment from the average:

$$FS = f(\frac{E_I}{N}, \boldsymbol{\alpha}_g, \boldsymbol{\beta}_g, \boldsymbol{\alpha}_{s_1}, \boldsymbol{\beta}_{s_1}, \ldots, \boldsymbol{\beta}_{s_4}) = f(\boldsymbol{x}) \quad (14)$$

where N is the number of visible vertices and $\boldsymbol{x}$ regroups the features from which the FS is computed.

The function $f(\cdot)$ is learned by a Classification Support Vector Machine [13]:

$$FS = \sum_{i=1}^{N_s} K(\boldsymbol{x}_i, \boldsymbol{x}) + b \quad (15)$$

where N_s is the number of support vectors and $\boldsymbol{x}_i$ is the ith support vector. The good fits are given the value $FS = 1$ and the poor fits $FS = -1$. Half of the data were used for training and the other half for validation. The SVM kernel, $K(\cdot)$ is Gaussian, and we chose an asymmetric cost function which penalizes positive deviations. This constraint limits the number of poor fits classified as correct fits (false positives) and ensures that the lower identification performance of the poor fits will not pollute the better identification performance of the accurate fits. Table 2 shows classification results over the training and validation sets.

Figure 3. Examples of model fitting. Top row: Initial parameters, Middle row: Results of fitting, rendered on top of the input images, Bottom row: Input images. The fifth column is an example of a poor fit.

5 Identification

Given a set of gallery images of faces, the identification system stores the model coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ computed by the fitting algorithm (Figure 1). In order to identify a person on a probe image, the fitting algorithm computes model coefficients, and uses the similarity measure (13) to find the nearest neighbour within the gallery.

The gallery is composed of a single image for each of the 68 individuals. All gallery images have the same illumination direction, which is close to the camera axis. The probe set is composed of $68 \cdot (3 \cdot 22 - 1) = 4420$ images of all 68 individuals at 3 poses and illuminated from 22 directions, but without the gallery images. We use the *closed universe* model for evaluating the performance, so each of the 68 persons in the probe set can be found in the gallery. None of the individuals shown in the gallery or probe images is in the database of 3D scans.

5.1 Identification Confidence

In this section, we demonstrate that the fitting score is a good measure of the identification confidence. Figure 4 shows the identification results with respect to the FS for a gallery of side views ($\approx 45°$). We divided the probe images into 8 bins of different FS and computed the percentage of correct identification for each of these bins (histogram in Figure 4). There is a strong correlation between fitting score and identification performance, indicating that FS is a good measure of identification confidence.

5.2 Identification Results

Figure 5 is the cumulative version of Figure 4. It shows that from the images with acceptable fits ($FS > 0$), which are 80 % of the images, 92.8 % are correctly identified, despite the wide pose and illumination variations. On the whole set of images, 82.6 % are correctly identified.

Table 3 presents identification results based on galleries composed of front, side and profile views, respectively. The table indicates that the best generalization performances are obtained with a side-view gallery. As expected, the poorest performances are obtained by a profile-view.

6 Conclusions

Given the large variations in illuminations and changes in viewpoint from front to profile, the performance of our algorithm seems promising. For further evaluation, the method needs to be applied to a larger dataset, since our current results are based on 68 individuals only.

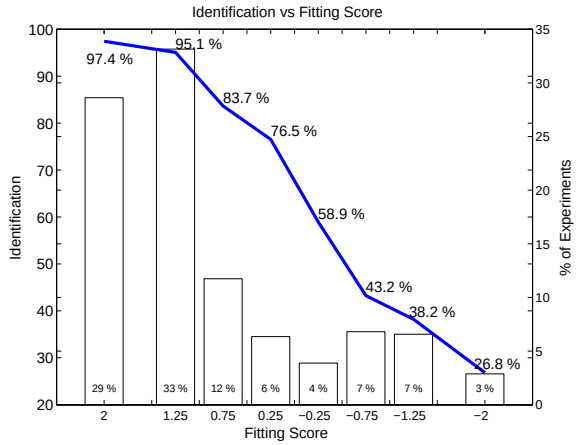

Figure 4. Identification results as a function of the fitting score.

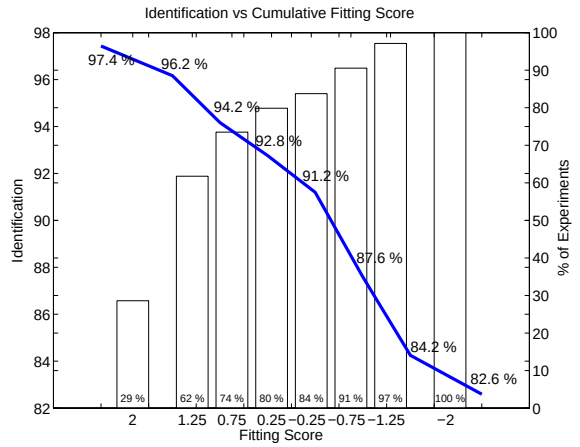

Figure 5. Cumulative identification results as a function of the fitting score.

The correlation between fitting score and identification performance provides us with a useful measure of confidence. The identification performance on those images that were rated with a high fitting score was 97.4%. This indicates that the model parameters of shape and texture are an appropriate representation of the identity of a face. It also suggests that higher identification performance on the entire set can be achieved by increasing the reliability of the fitting algorithm.

Currently, the system requires some manual interaction for initializing the fitting process, but we are developing methods to alleviate this, and speed up the fitting process.

Acknowledgements: This work was supported by the European Research Office of the US Army under contract No. N68171-01-C-9000. The database of laser scans was recorded at Max-Planck-Institute for Biological Cybernetics, Tübingen.

		probe view					
		front		side		profile	
front view gallery	mean	98	94	94	85	69	65
	std	2.0	6.3	8.0	20.7	18.3	18.2
side view gallery	mean	95	89	98	90	78	70
	std	3.1	6.4	3.0	9.2	16.9	18.9
profile view gallery	mean	76	71	81	71	89	84
	std	7.2	9.2	10.4	12.2	12.5	16.4

Table 3. Mean identification percentage and their standard deviations averaged over all lighting conditions for front, side and profile view galleries. The first number of each condition is based on the good fits only ($FS > 0$), the second on the entire probe set.

References

[1] J. Bergen and R. Hingorani. Hierarchical motion-based frame rate conversion. Technical report, David Sarnoff Research Center Princeton NJ 08540, 1990.

[2] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *Computer Graphics Proceedings SIGGRAPH'99*, pages 187–194, Los Angeles, 1999.

[3] T. F. Cootes, K. Walker, and C. J. Taylor. View-based active appearance models. In *Int. Conf. on Autom. Face and Gesture Recognition*, pages 227–232, Grenoble, 2000.

[4] J. Foley and A. v. Dam. *Fundamentals of interactive computer graphics*. The systems programming series. Addison-Wesley, Reading, Ma, 1984.

[5] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.

[6] P. Hallinan. *A deformable model for the recognition of human faces under arbitrary illumination*. PhD thesis, Harvard University, Cambridge, Massachusetts, 1995.

[7] R. Haralick and L. Shapiro. *Computer and robot vision*, volume 2. Addison-Wesley, Reading, Ma, 1992.

[8] S. Haykin. *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 1998.

[9] M. Jones and T. Poggio. Multidimensional morphable models: A framework for representing and matching object classes. *Int. Journal of Comp. Vision*, 29(2):107–131, 1998.

[10] A. Lanitis, C. Taylor, and T. Cootes. Automatic face identification system using flexible appearance models. *Image and Vision Computing*, 13(5):393–401, 1995.

[11] H. Moon and P. J. Phillips. Computational and performance aspects of pca-based face-recognition algorithms. *Perception*, 30:303–321, 2001.

[12] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression (PIE) database of human faces. Technical Report CMU-RI-TR-01-02, The Robotics Institute, Carnegie Mellon University, January 2001.

[13] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 1995.

[14] T. Vetter and T. Poggio. Linear object classes and image synthesis from a single example image. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(7):733–742, 1997.

[15] W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips. Face recognition: A literature survey. 2000. submitted.