Face Reconstruction from a Small Number of Feature Points*

Bon-Woo Hwang†, Volker Blanz‡, Thomas Vetter‡ and Seong-Whan Lee†

†Center for Artificial Vision Research, Korea University
Anam-dong, Seongbuk-ku, Seoul 136-701, Korea
{bwhwang, hhsong, swlee}@image.korea.ac.kr

‡Max-Planck-Institute for Biological Cybernetics
Spemannstr. 38, 72076 Tuebingen, Germany
{volker.blanz, thomas.vetter}@tuebingen.mpg.de

Abstract

This paper proposes a method for face reconstruction that makes use of only a small set of feature points. Faces can be modeled by forming linear combinations of prototypes of shape and texture information. With the shape and texture information at the feature points alone, we can achieve only an approximation to the deformation required. In such an under-determined condition, we find an optimal solution using a simple least square minimization method. As experimental results, we show well-reconstructed 2D faces even from a small number of feature points.

1. Introduction

It is difficult for traditional bottom-up, generic approaches to reconstruct the whole image of an object from parts of its image or to restore the missing space information due to noise, occlusion by other objects, or shadow caused by illumination effects. In contrast to such approaches, top-down, object-class-specific and model-based approaches are highly tolerant to sensor noise and incompleteness of input image information[2]. Hence, the top-down approaches to the interpretation of images of variable objects are now attracting considerable interest[1-4, 6]. Kruger et al. proposed a system for the automatic determination of the position, size and pose of a human head, which is modeled by a labeled graph[3]. The nodes of the graph refer to feature points of a head. However, the location and texture information are not utilized for face reconstruction. Lanitis et al. implemented a face reconstruction using a 'Flexible Model'[4]. About 150 feature points inside a face and its contour were used for reconstruction. The system performs well, but it requires more than 100 feature points for shape reconstruction, and it also requires full texture information.

In this paper, we propose an efficient face reconstruction method from a small set of feature points. Our approach is based on the 2D shapes and textures of a data set of faces. Shape is coded by a dense deformation field between a reference face image and each individual face image. Given only a small number of feature points on a novel face, we use the database to reconstruct its full shape and texture. It is a combination of stored shapes and textures that best matches the positions and gray values of the feature points. We anticipate that the proposed method will play an important role in reconstructing the whole information of an object out of information reduced for compressing or partially damaged due to occlusion or noise. In Section 2 and 3, we describe a 2D face model where shape and texture are treated separately, and a method for finding coefficients for face reconstruction, respectively. Experimental results for face reconstruction are given in Section 4. Finally, in Section 5, we present conclusive remarks and discuss some future work.

2. 2D face model

On the assumption that the correspondence on the face images has already been established[1], the 2D shape of a face is coded as the deformation field from a reference image that serves as the origin of our space. The texture is coded as the intensity map of the image which results from mapping the face onto the reference face[6].

Let $S(x)$ be the displacement of point x, or the position of the point in the face that corresponds to point x in the reference face. Let $T(x)$ be the gray value of the

*This research was supported by Creative Research Initiatives of the Ministry of Science and Technology, Korea.

point in the face that corresponds to point x in the reference face. With shape and texture information separated from the face image, we fit a multivariate normal distribution to our data set of faces according to the average of shape $\bar{S}$ and that of texture $\bar{T}$ and covariance matrices C_S and C_T computed over shape and texture differences $\tilde{S} = S - \bar{S}$ and $\tilde{T} = T - \bar{T}$. By Principal Component Analysis(PCA), a basis transformation is performed to an orthogonal coordinate system formed by eigenvectors s_i and t_i of the covariance matrices on our data set of m faces.

$$S = \bar{S} + \sum_{i=1}^{m-1} \alpha_i s_i,\ T = \bar{T} + \sum_{i=1}^{m-1} \beta_i t_i, \tag{1}$$

where $\vec{\alpha},\ \vec{\beta}\ \epsilon\ \Re^{m-1}$. The probability for coefficients $\vec{\alpha}$ is defined as

$$p(\vec{\alpha}) \sim exp\left[-\frac{1}{2}\sum_{i=1}^{m-1}(\frac{\alpha_i}{\sigma_i})^2\right], \tag{2}$$

with σ_i^2 being the eigenvalues of the shape covariance matrix C_S. Likewise, the probability $p(\vec{\beta})$ can be computed.

3. Face reconstruction

In this section, we describe a method for finding coefficients for face reconstruction. First, we define an energy function as the sum of normalized coefficients and set a condition for minimizing the energy function. Then, we solve this problem by a least square minimization.

3.1. Problem definition

Since there are shape and texture elements only for feature points, we achieve an approximation to the deformation required. Our goal is to find an optimal solution in such an underdetermined condition. We define an energy function as the sum of the normalized coefficients. We also set a condition that the given shape or texture information at the feature points must be reconstructed perfectly. The energy function, $E(\alpha)$, describes the degree of deformation from the the reference face. The problem(Equation 3) is to find α which minimizes the energy function, $E(\alpha)$, which is given as:

$$\alpha^* = arg\min_{\alpha} E(\alpha), \tag{3}$$

with the energy function,

$$E(\alpha) = \sum_{i=1}^{m-1}(\frac{\alpha_i}{\sigma_i})^2. \tag{4}$$

under the condition,

$$\tilde{S}(x_j) = \sum_{i=1}^{m-1} \alpha_i s_i(x_j),\ (j = 1,\ \cdots,\ n), \tag{5}$$

where $x_1,\ \cdots,\ x_n$ are the selected feature points. Since we select only a small number of feature points, n is much smaller than $m-1$.

3.2. Solution by least square minimization

According to Equation 3∼5, we can solve this problem using general quadratic programming. In order to make this problem simpler, we reduce it to a least square problem. Equation 5 is equivalent to the following:

$$\begin{pmatrix} s_1(x_1) & \cdots & s_{m-1}(x_1) \\ \vdots & \ddots & \vdots \\ s_1(x_n) & \cdots & s_{m-1}(x_n) \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_{m-1} \end{pmatrix} = \begin{pmatrix} \tilde{S}(x_1) \\ \vdots \\ \tilde{S}(x_n) \end{pmatrix}. \tag{6}$$

To exploit the inherent orthogonal nature of the problem, we rewrite Equation 6 as:

$$\mathbf{S}\ \alpha' = \tilde{\mathbf{S}}, \tag{7}$$

where

$$\mathbf{S} = \begin{pmatrix} \sigma_1 s_1(x_1) & \cdots & \sigma_{m-1} s_{m-1}(x_1) \\ \vdots & \ddots & \vdots \\ \sigma_1 s_1(x_n) & \cdots & \sigma_{m-1} s_{m-1}(x_n) \end{pmatrix},$$

$$\alpha' = (\frac{\alpha_1}{\sigma_1}, \cdots, \frac{\alpha_{m-1}}{\sigma_{m-1}})^T,$$

$$\tilde{\mathbf{S}} = (\tilde{S}(x_1), \cdots, \tilde{S}(x_n))^T, \tag{8}$$

and the row vectors of $\mathbf{S}$ are assumed to be linearly independent. α' can be computed by

$$\alpha' = \mathbf{S}^+\ \tilde{\mathbf{S}}, \tag{9}$$

where $\mathbf{S}^+$ is the pseudoinverse of the matrix $\mathbf{S}$, and can be obtained easily using a singular value decomposition as follows[5].

Supposing the singular value decomposition of $\mathbf{S}$ is

$$\mathbf{S} = U\ W\ V^T, \tag{10}$$

the pseudoinverse of $\mathbf{S}$ is

$$\mathbf{S}^+ = V\ W^+\ U^T. \tag{11}$$

The columns of U are eigenvectors of $\mathbf{S}\mathbf{S}^T$, and the columns of V are eigenvectors of $\mathbf{S}^T\mathbf{S}$. The main diagonals of W are filled with the square roots of the nonzero eigenvalues of both. In W^+, all nonzero elements of W are replaced by their reciprocals.

Figure 1 represents the process described above in 2D-α' space. This is the case that, as we previously assumed, the row vectors of $\mathbf{S}$ are linearly independent, and the number of the bases m-1 and the feature point n are 2 and 1, respectively. Circular contour lines designate two-dimensional probability plots of Gaussian, $P(\alpha)$. The solid line represents the condition given in Equation 5.

The point at which the energy function is minimized can be obtained by finding the point which has the minimum distance from the origin.

Using Equations 1, 8 and we obtain

$$S = \bar{S} + \sum_{i=1}^{m-1} \alpha'_i \sigma_i s_i. \tag{12}$$

By using Equation 12, we can get correspondence of all points. Similarly, we can construct full texture information T.

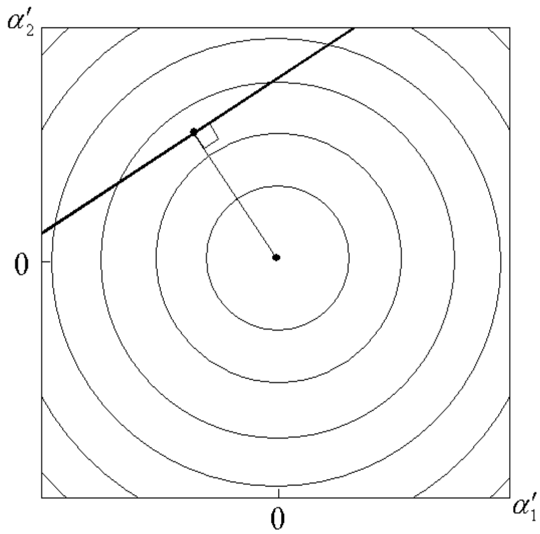

Figure 1. The Example of least square minimization in 2D-α' space

We previously made the assumption that the row vectors of $\mathbf{S}$ in Equation 8 are linearly independent. Otherwise, Equation 5 may not be satisfied. In other words, the correspondence obtained from Equation 12 may be inaccurate not only between the feature points but also at the feature points. Therefore, for our purpose of effectively reconstructing a face image from a few feature points, selecting the feature points that are linearly dependent would not be appropriate. However, this is unlikely to happen.

4. Experimental results

For testing the proposed method, we used 200 two-dimensional images of human faces that were rendered from a database of three-dimensional head models recorded with a laser scanner($Cyberware^{TM}$)[1, 6]. The face images had been collected for psychophysical experiments from males and females between twenty and forty years old. They wear no glasses and earrings. Males must have no beard. The resolution of the images was 256 by 256 pixels and the color images were converted to 8-bit gray level images. The images were generated under controlled illumination conditions and the hair of the heads was removed completely from the images. PCA is performed on a random subset of 100 face images. The other 100 images are used to test the algorithm.

4.1. Selection of feature points

For face reconstruction from feature points, we first set the number and location of the feature points to be used in the reference face. The feature points from all faces can be automatically extracted by using the given correspondence once the feature points have been selected in the reference face. In Figure 2, the white cross points represent the feature points selected for shape reconstruction. For reconstruction of texture information, additional points are selected, and they are represented by black cross points in the figure. In order to reduce errors caused by noise(e.g. salt and pepper noise), the mid value of the p by p mask that runs on each point is obtained and collected as texture information. In our experiments, 22 feature points are chosen for shape reconstruction and 3 more for texture reconstruction, and A 3 by 3 mask is used for error reduction.

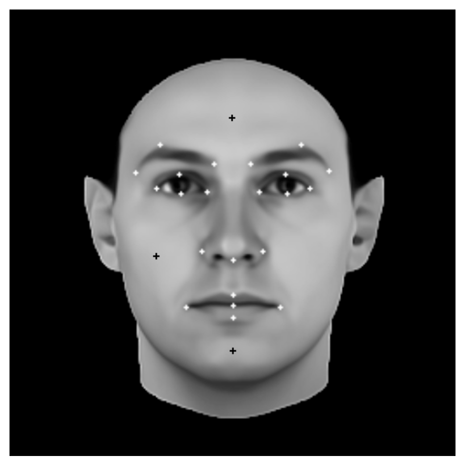

Figure 2. The selected feature points

4.2. Reconstruction of shape and texture

For testing the algorithm, we use the fact that we already have the correct correspondence for the test images, and assume that we knew only the location of the feature points.

Using the correspondence we can automatically extract the feature points from all faces, once the feature points have been selected in the reference faces.

As mentioned before, 2D-shape and texture of face images can be treated separately. Therefore, A face image can be constructed by combining shape and texture information after reconstruction of both information. In Figure 3, the shape information of the face images is reconstructed from correspondence at the feature points. Instead of texture extraction, we used the standard texture of the reference face image. The images on the left are the original face images and those on the right are the face images reconstructed by the proposed method. Only the images of the inside of the face region are presented in the figure for the convenience of comparison of the reconstructed face images with the original ones. Figure 4 shows face images reconstructed both from shape and texture information at feature points. In contrast to the reconstructed face images in Figure 3, where texture is the same as that of the reference face image, it can be easily noticed that the brightness of the skin is successfully reconstructed.

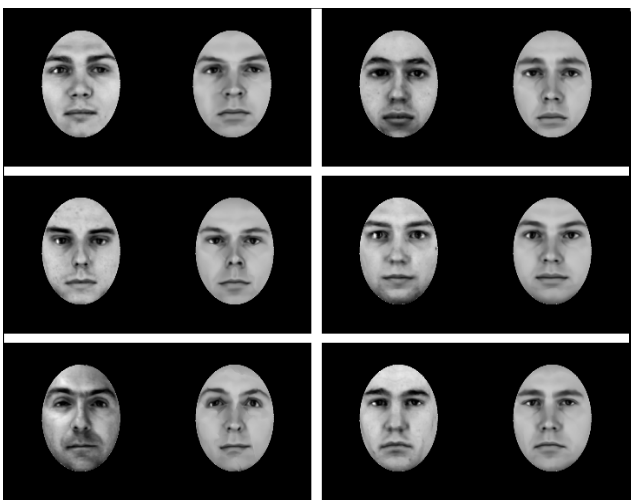

Figure 3. Examples of shape reconstruction (left: originals, right: reconstructions)

5. Conclusions and future work

In this paper, we have proposed an efficient face reconstruction method from a small set of feature points. The proposed method uses a strategy that minimizes face deformation, provided that the shape and texture of feature points are perfectly reconstructed. As experimental results, well-reconstructed 2D face images similar to original ones are obtained. However, the feature points are selected heuristically. They need to be automatically chosen linearly independent for the proposed method. Therefore, an elegant algorithm of selecting feature points is required for efficient reconstruction. The proposed method is expected to be useful for reconstruction of information reduced or partially damaged.

Figure 4. Examples of shape and texture reconstruction(left: originals, right: reconstructions)

Acknowledgement

The authors would like to thank Max-Planck-Institute for providing the MPI Face Database.

References

[1] V. Blanz and T. Vetter. A morphable model for the synthesis of 3D faces. *Proc. of SIGGRAPH'99, Los Angeles*, pages 187–194, August 1999.

[2] M. J. Jones, P. Sinha, T. Vetter, and T. Poggio. Top-down learning of low-level vision tasks[brief communication]. *Current Biology*, 7:991–994, 1997.

[3] N. Kruger, M. Potzsch, and C. von der Malsburg. Determination of face position and pose with a learned representation based on labelled graphs. *Image and Vision Computing*, 15:665–673, 1997.

[4] A. Lanitis, C. J. Taylor, and T. F. Cootes. Automatic interpretation and coding of face images using flexible models. *IEEE Trans. on Pattern Anaysis and Machine Intelligence*, 19(7):743–756, 1997.

[5] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipies in C.* Cambridge University Press, Port Chester, NY, 1992.

[6] T. Vetter and N. E. Troje. Separation of texture and shape in images of faces for image coding and synthesis. *Journal of the Optical Society of America A*, 14(9):2152–2161, 1997.