2011

2D Face Database Diversification Based on 3D Face Modeling

Qun Wang
Old Dominion University

Jiang Li
Old Dominion University, jli@odu.edu

Vijayan K. Asari
University of Dayton

Mohammad A. Karim
Old Dominion University

Manuel Filipe Costa (Ed.)

Original Publication Citation

Wang, Q., Li, J., Asari, V. K., & Karim, M. A. (2011) 2D face database diversification based on 3D face modeling. In M. F. Costa (Ed.), *International Conference on Applications of Optics and Photonics, Proceedings of SPIE Vol. 8001* (80010M). SPIE of Bellingham, WA. https://doi.org/10.1117/12.894605

2D Face Database Diversification Based on 3D Face Modeling

Qun Wang[a], Jiang Li[a], Vijayan K. Asari[b], Mohammad A. Karim[a]
[a]Dept. of Electrical and Computer Engineering, Old Dominion University, VA, USA 23508;
[b]Dept. of Electrical and Computer Engineering, University of Dayton, OH, USA 45469

ABSTRACT

Pose and illumination are identified as major problems in 2D face recognition (FR). It has been theoretically proven that the more diversified instances in the training phase, the more accurate and adaptable the FR system appears to be. Based on this common awareness, researchers have developed a large number of photographic face databases to meet the demand for data training purposes. In this paper, we propose a novel scheme for 2D face database diversification based on 3D face modeling and computer graphics techniques, which supplies augmented variances of pose and illumination. Based on the existing samples from identical individuals of the database, a synthesized 3D face model is employed to create composited 2D scenarios with extra light and pose variations. The new model is based on a 3D Morphable Model (3DMM) and genetic type of optimization algorithm. The experimental results show that the complemented instances obviously increase diversification of the existing database.

Keywords: 3D face modeling, 3D morphable model, differential evolution, face recognition

1. INTRODUCTION

Pose and illumination are identified as major problems in 2D face recognition (FR) [1][2]. It has been theoretically proven that the more diversified instances in the training phase, the more accurate and adaptable the FR system appears to be [3]. In recent decades, researchers have developed various photographic face databases to meet the demands for face training (Figure 1). However, these databases can only supply images with limited variations.

In this paper, we propose an approach to diversify the photographic image database by using the representation and synthesis of 3D model of human faces. The new 3D face model associated with computer graphics techniques is able to accomplish the goal of providing augmented variances of pose and illumination.

The remaining sections are organized as follows: The concept of morphable model is introduced in Section 2. Model matching method based on Differential Evolution is presented in Section 3. The final experimental results are outlined in Section 5 and Conclusions in Section 6.

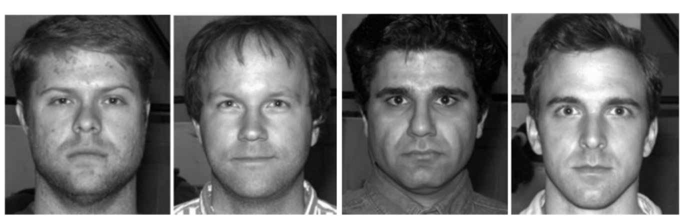


Figure 1. Face images taken from different subjects

(Courtesy of Yale Face Database B)

International Conference on Applications of Optics and Photonics, edited by Manuel F. M. Costa,
Proc. of SPIE Vol. 8001, 80010M · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.894605

2. 3D MORPHABLE MODEL

2.1 3D Face Modeling

3D face modeling has been one of the greatest challenges for researchers in computer graphics for many years. 3D Morphable Model (3DMM) proposed by T. Vetter [5][6] is a realistic 3D face modeling method which could be synthesized automatically by linear combination of exemplar faces. One of the applications of 3DMM in face recognition is to create the 3D face model of an individual from given 2D images [4][5].

The reconstruction procedure is regarded as conducting iterations of the analysis-by-synthesis process, which are driven by fitting the 3D model to 2D images. Meanwhile, the parameters with respect to 3D environment such as focal length of the camera, illumination and color contrast, can also be modeled explicitly and estimated automatically.

The morphable model has several advantages. Unlike other models, such as Shape-From-Shading (SFS) [7], it has no restriction on the requirement of illumination or reflectance functions though it has additional computational complexity.

As valuable as it is, morphable model, therefore, could be utilized to expand the spectrum and create versatile variations for the original photographic database [8].

2.2 Model Construction

The prototypical 3D faces are acquired by 3D laser scanners, whose range and texture data are digitalized with high precision. Preprocessed through registration and texture extraction, each face is represented in the form of a shape vector and a texture vector as:

$$S = (X_1, Y_1, Z_1, X_2, ..., Y_n, Z_n)^T \quad (1)$$

$$T = (R_1, G_1, B_1, R_2, ..., G_n, B_n)^T \quad (2)$$

where n is the number of vertexes on the 3D face and (B_j G_j R_j) are the corresponding R, G, B color values of the vertex (X_j Y_j Z_j). Therefore, a morphable model can be generated by using the linear combination of shape vectors S_i and texture vectors T_i of 3D training faces as [6]:

$$S_{model} = \sum_{i=1}^{m} a_i S_i \qquad T_{model} = \sum_{i=1}^{m} b_i T_i \quad (3)$$

$$\sum_{i}^{m} a_i = \sum_{i}^{m} b_i = 1$$

in which m is the number of training faces, Si and Ti are shape and texture of training faces and a_i and b_i are their corresponding weights contributed to the new face with $0 < a, b < 1$.

In the practical consideration of computational effectiveness, a common technique as PCA (Principal Component Analysis) is employed to reduce the high dimensionality of 3D face data without the loss of potential face information. In particular, PCA performs a transformation of the original cloud data to an orthogonal coordinate system formed by the eigenvectors s_i and t_i of the covariance matrices.

$$\mathbf{S}_{model} = S_{mean} + \sum_{i=1}^{m-1} \alpha_i \mathbf{s}_i \qquad T_{model} = T_{mean} + \sum_{i=1}^{m-1} \beta_i T_i \quad (4)$$

where S_{mean} and T_{mean} are the average shape and texture vectors. S_i and T_i are principal components. $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$ and $\beta = (\beta_1, \beta_2, ..., \beta_n)$ are shape and texture combination coefficients, and α and β obeys Gaussian distribution as:

$$p(\boldsymbol{\alpha}) = \exp\left(-\frac{1}{2}\sum_{i=1}^{m-1}\left(\frac{\alpha_i}{\sigma_{S,i}}\right)^2\right) \qquad p(\beta) = \exp\left(-\frac{1}{2}\sum_{i=1}^{m-1}\left(\frac{\beta_i}{\sigma_{T,i}}\right)^2\right) \tag{5}$$

2.3 Model Matching

Matching the 3D face morphable model to the given face images is a process of model parameter estimation, in which a bunch of coefficients involved are required to be determined. For example, camera and illumination model is adopted in the projection of the 3D face model into the image plane since 3D face model and 2D input facial images cannot be measured directly. Aiming at retrieving a 3D face the closest projective image to the input facial image, the error function between 3D model projective image I_{mod} and input image I_{input} is described as:

$$E_I = \sum_{x,y} \left\| I_{input}(x,y) - I_{\mathrm{mod}}(x,y) \right\|^2 \tag{6}$$

In order to create a realistic 2D output which is close enough to the target face image, we make use of the perspective projection and Phong illumination model in the rendering process. Given the k^{th} vertex at (X,Y,Z) with texture value (R,G,B), the perspective projection on the image plane is represented as:

$$I_k(x,y) = \left(I_{r,k}(x,y), I_{g,k}(x,y), I_{b,k}(x,y)\right)^T \tag{7}$$

where $I_{c,k}(x,y)$ is computed under the Phong illumination model as:

$$I_{c,k}(x,y) = R\left(I_{a,c} + I_{dir,c}(L \cdot N)\right) + K_s\, I_{dir,c}(F \cdot V)^n \tag{8}$$

$I_{a,c}$ and $I_{dir,c}$ are separately intensity of ambient light and direct light of the c^{th} color component. K_s is the reflectance, L, N, F, V are light direction, normal, reflective direction and direction of viewer respectively and n is the mirror reflectance index.

3. GLOBAL OPTIMIZATION

The morphable model provides a new approach in solving difficult problems of facial recognition research with extreme illumination and pose variations. However, one of the issues lies in the process of image matching that performs error evaluation in the pixel-level measurement. This involves algorithms of image matching and a large-scale optimization. In our current research, stochastic gradient descent [6] method is used to evaluate the residual and global error as well as objective function optimization.

In 3DMM, the estimation process of fitting shape and texture information into 2D images is considered as the issue of searching for the global minimum in a high dimensional feature space, in which optimization is apt to have local convergence. On the other hand, Differential Evaluation (DE) appears to be robust against stagnation in local minima and sensitiveness to initial values in face reconstruction. Considering its successful performance, we tentatively introduce DE to tackle the problem in 3D-2D matching.

3.1 Differential Evolution (DE)

Differential Evolution (DE) is a "parallel direct search method" [9], which was first proposed by Storn and Price in 1995 [10]. It is characterized as a stochastic and population-based optimization that is simple and effective for implementation. DE repeatedly processes through operations which are, in turn, as "mutation, crossover and selection" until an optimal solution to the objective function $f(x)$ is reached (Figure 2).

The classic version of DE is defined as follows. Suppose we have N D-dimensional parameter vectors

$$x_{i,G} = \left[x_1, x_2, ..., x_D\right]^T, i = 1, 2, ..., N \tag{9}$$

representing the population for generation *G*. The algorithm starts by randomly initializing the vector populations with, as the author suggested, a uniform probability distribution [9]. We use a different distribution in our experiment due to the special feature of 3DMM, which we will present later in section 4.

3.1.1 Mutation

For each individual x_i, a corresponding mutation vector v_i is produced according to the equation:

$$v_{i,G} = x_{r1,G} + F * \left(x_{r2,G} - x_{r3,G}\right) \tag{10}$$

in which random index $r1, r2, r3 \in \{1, 2, ..., N\}$ and $r1 \neq r2 \neq r3$. *F* is a real amplifier designed to control the offset of $v_{i,G}$ to $x_{r1,G}$ by scaling the differential variation $\left(x_{r2,G} - x_{r3,G}\right)$.

3.1.2 Crossover

Trial vectors are introduced in the phase of crossover to expand the range of global search. It is defined in the form as:

$$u_{i,G} = \left(u_{1i,G}, u_{2i,G}, ..., u_{Di,G}\right) \tag{11}$$

in which

$$u_{ji,G} = \begin{cases} v_{ji,G}, if\left(randb(j) \le CR\right) \vee j = I_{br} \\ x_{ji,G}, if\left(randb(j) > CR\right) \wedge j \neq I_{br} \end{cases} \qquad j = 1, 2, ..., D. \tag{12}$$

In equation (9), $randb(\cdot)$ is a random generator with uniform distribution. I_{br} is an integer randomly chosen from *{1,2,...,D}*, which prevents $x_{i,G}$ from being equal to $u_{i,G}$.

3.1.3 Selection

DE utilizes pair-wise comparison between $u_{i,G}$ and $x_{i,G}$ to survive the vectors with fewer objectives function values to the next generation.

$$x_{i,G+1} = \begin{cases} u_{i,G}, if\left(f\left(u_{i,G}\right) < f\left(x_{i,G}\right)\right) \\ x_{i,G}, otherwise \end{cases} \qquad i = 1, 2, ..., N \tag{13}$$

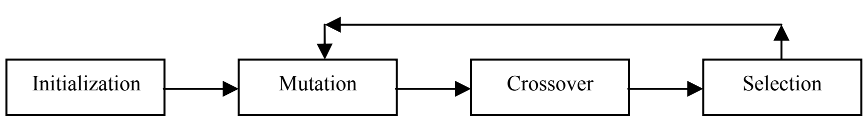


Figure 2. Classic Differential Evolution Procedure

4. EXPERIMENTAL RESULTS

4.1 3D Face Modeling

To derive a morphable model for 3D shape and texture, we use the database from the 3D Basel Face Model (BFM) [8]. The database is created from 100 male faces and 100 female faces, each of which is keeping neutral expression, without makeup, accessories and glasses. The registered 3D faces are parameterized as triangular meshes with 53490 vertices and shared topology [8]. Figure 3 shows the mean face of the 200 faces in the database, which is represented as S_{mean} and T_{mean} in (4).

Figure 3. 3D mean face from BMF database, computed from 100 male faces and 100 female faces

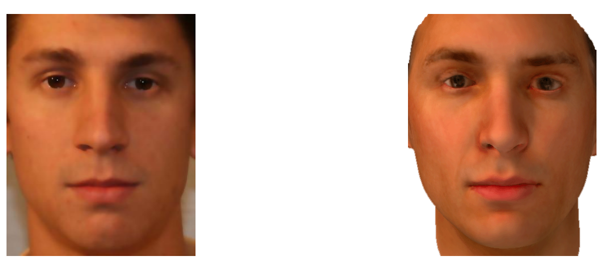
4.2 3D Face Model Training

As described in section 3, 3D morphable model equipped with DE intuitively leads to our framework for 3D model reconstruction. In our experiment, DE populations are initialized by Gaussian distribution instead of uniform distribution. The reason for this is, according to the equation (5), shape and texture combination coefficients, α and β, both obey Gaussian distribution.

BFM database provides 200 principal components for both shape and texture representations. Due to the consideration of computational efficiency, we only employ the first 70 components of S_i and T_i for face training. Even then, final results still show that these components are competent for sound outcome.

4.3 Simulation Results

In this section, we present experimental results covering from lighting and pose variations to facial appearance variations. We make use of partial images from the ODU-VL face database.



(a) (b)

Figure 4. original image and synthesized image, (a) original image; (b) synthesized image

4.3.1 Lighting Variations

Once the 3D model is reconstructed, it is ready to re-render the 3D scene by using different lighting variations. Figure 5 shows composited scenarios dominated by light sources in different positions. Given constant distance from the light to the object, we change the light source positions by varying azimuth and elevation angles of the light source.

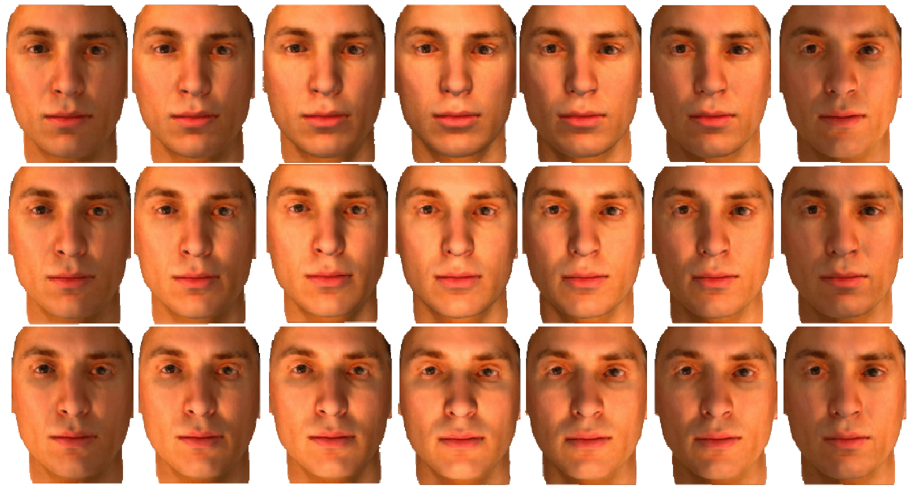


Figure 5. 3D faces rendered under different illumination conditions; azimuth and elevation angles ranging from -60 to 60 and -30 to 30 respectively.

4.3.2 Pose Variations

Figure 6 shows composited scenarios rendered by various pose conditions. We diversify the face orientation by assigning different azimuth and elevation angles to the transform matrix of rotation.

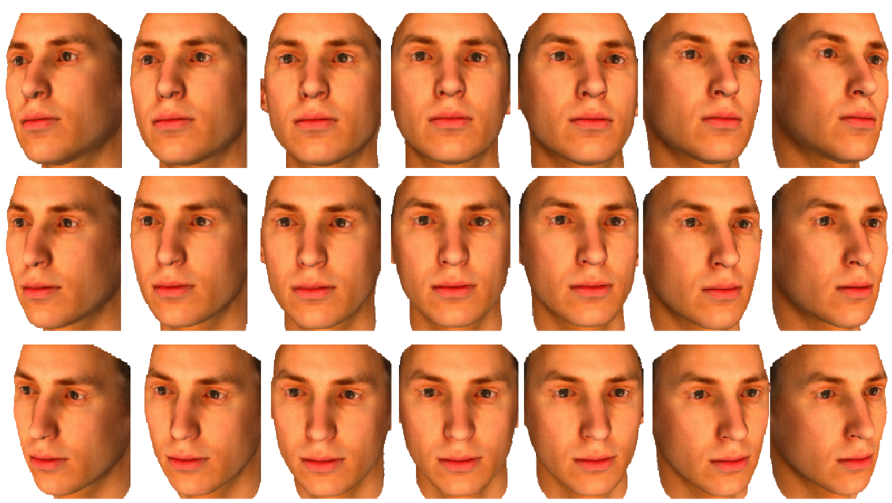


Figure 6. 3D faces rendered under different pose conditions; azimuth and elevation angles ranging from -30 to 30 and -10 to 10 respectively.

5. CONCLUSIONS

Pose and illumination are regarded as major problems for FR systems. In practice, photographic face databases are not able to provide as many variances as researchers expect. In this paper, we propose a novel scheme for 2D face database diversification, which supplies augmented variances of pose and illumination. The experimental results show that the complemented instances obviously vary representations of face images and diversify the database. Future research will be focused on the computation efficiency of DE optimization.

6. ACKNOWLEDGEMENT

The authors would like to express thanks to Dr. T, Vetter, Department of Computer Science, and the University of Basel, for providing the 3D Basel Face Model database for our experiments. The authors also acknowledge the individual whose image is used in our experiments.

REFERENCES

[1] Phillips, P. J. , Grother, P. , Micheals, R. , Blackburn, D. M. , Tabassi, E. , Bone, M. , "Face recognition vendor test 2002," Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 44 (2003)

[2] Adini, Y., Moses, Y. and Ullman, S., "Face recognition: the problem of compensating for changes in illumination direction," Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19(7), 721-732(1997).

[3] Zhao, W., Chellappa, R., Phillips, P. J. and Rosenfeld, A., "Face recognition: A literature survey," ACM Computing Surveys, 35(4),399–458 (2003).

[4] Blanz, V. and Vetter, T., "A Morphable Model For The Synthesis Of 3D Faces," SIGGRAPH '99: Proc. of the 26th annual conference on Computer graphics and interactive techniques, 187-194 (1999).

[5] Parke, F. I., "A Parametric Model of Human Faces", PhD thesis, Salt Lake City: University of Utah (1974).

[6] Vetter, T. and Poggio, T., "Linear object classes and image synthesis from a single example image," Pattern Analysis and Machine Intelligence, IEEE Transactions on, 19(7), 733-742 (1997).

[7] Atick, Joseph J., Griffin, Paul A. and Redlich, A. Norman, "statistical Approach to Shape from Shading: Reconstruction of Three-Dimensional Face Surfaces from Single Two-Dimensional Images", Neural Computation, 8(6), 1321-1340 (1996).

[8] Paysan, P., Knothe, R., Amberg, B., Romdhani, S., and Vetter, T., "A 3D Face Model for Pose and Illumination Invariant Face Recognition," Advanced Video and Signal Based Surveillance, AVSS '09. Sixth IEEE International Conference on, 296-301, (2009).

[9] Storn, R. and Price, K., "Differential Evolution - a Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces," Technical Report TR-95-012, ICSI, (1995).

[10] Storn, R. and Price, K., "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces", Journal of Global Optimization, 11(4), 341-359 (1997).