https://doi.org/10.17576/apjitm-2014-0301-02

http://www.ftsm.ukm.my/apjitm
Asia-Pacific Journal of Information Technology and Multimedia
Jurnal Teknologi Maklumat dan Multimedia Asia-Pasifik
Vol. 3 No. 1, June 2014 : 15 - 21
e-ISSN: 2289-2192

RECONSTRUCTION OF 3D FACES BY SHAPE ESTIMATION AND TEXTURE INTERPOLATION

ASHRAF Y. A. MAGHARI
IBRAHIM VENKAT
BAHARI BELATON

ABSTRACT

This paper aims to address the ill-posed problem of reconstructing 3D faces from single 2D face images. An extended Tikhonov regularization method is connected with the standard 3D morphable model in order to reconstruct the 3D face shapes from a small set of 2D facial points. Further, by interpolating the input 2D texture with the model texture and warping the interpolated texture to the reconstructed face shapes, 3D face reconstruction is achieved. For the texture warping, the 2D face deformation has been learned from the model texture using a set of facial landmarks. Our experimental results justify the robustness of the proposed approach with respect to the reconstruction of realistic 3D face shapes.
Keywords: 3D face reconstruction, Tikhonov regularization, texture registration, PCA, TPS.

INTRODUCTION

The objective of 3D facial reconstruction systems is to recover the three-dimensional shape of individuals from their 2D pictures or video sequences. The problem of 3D facial modeling remains as a partially solved problem in the field of computer vision in terms of the accuracy and speed of reconstruction algorithms. This paper presents an approach for reconstructing the 3D face of an individual given the 2D face image where prior knowledge has been subsumed. The prior knowledge can be considered as general information about the shape properties of 3D face shape objects. These properties are learned from a set of 3D face shapes. With the help of the prior knowledge, the 3D face shape is estimated using a set of 2D control points, while the 2D texture is registered with the prior model and warped to the reconstructed 3D face shape.

The 3D shape learning model in this study relies on examples of 3D scans; this means that missing information can be inferred using correlation among the model shape vectors. For robust, plausible, and stable results, the regularization mechanism needs to find a tradeoff between fitting 3D shapes to the given 2D facial landmarks and producing plausible solutions in terms of prior knowledge (Blanz et al., 2004). The standard Tikhonov regularization method is used to estimate the model parameters by solving the inverse problem and preventing overfitting. However, the quality of the reconstructed face shapes is similar to the mean face shape (excessive smoothness); this leads to a loss of information about the reconstructed images (Jing et al., 2009). In this paper, Tikhonov regularization has been extended by replacing the identity matrix with the diagonal eigenvalue matrix. On the other hand, the Thin Plate Spline (TPS) technique is used to register the input images with the model texture. Then, the registered (interpolated) texture is warped to the reconstructed 3D shape face. The results have showed good 3D reconstructions and retained the real characteristics of the given 2D face images.

This paper is organized as follows: Related work is introduced in Section 2. In Section 3, modeling 3D face shapes is explained. Section 4 is devoted to the 3D shape model fitting to new faces, and Section 5 describes Texture Warping. Section 6 deals with Experiments and Discussion while the conclusion is reported in the last section, Section 7.

RELATED WORK

To provide more realistic face reconstructions than in other methods, example-based modeling could be utilized (Widanagamaachchi & Dharmaratne, 2008), (Levine & Yu, 2009). In this respect, example-based 3D face reconstruction methods have mainly two stages: the model building stage and the model fitting stage. In this paper, the PCA-based 3D face model is used for model building, and the regularized algorithm is used for model fitting. Some statistical modeling methods model both shape and texture separately using PCA (e. g., 3DMM (Blanz & Vetter, 1999)). Nevertheless, it has been suggested that shapes are more amenable to the PCA-based modeling than textures because textures vary more dramatically than shapes (Maghari, Liao & Belaton, 2012). In many situations, the 2D textures can be warped to the 3D geometry to generate the face textures (Jiang et al., 2005). Therefore, this study focuses on shape modeling where the reconstruction of 3D faces from single 2D images using shape models is relatively simple. One of the reconstruction methods that use prior knowledge to estimate the shape coefficients from a set of facial points is regularization (Jiang et al., 2005), (Blanz and Vetter, 2002). In (Jiang et al., 2005), the authors use a regularization equation that estimates the geometry coefficient in an iterative procedure. Alternatively, the regularization method has been presented in (Blanz & Vetter, 2002). The authors estimated the geometry coefficients directly, but they did not show any experimental results with respect to real 2D face images. Instead, experiments on test 3D face shapes have been shown.

MODELING 3D FACE SHAPES

The characteristic shape properties of the 3D face shapes are derived from a dataset of 3D scans. The 3D shapes are aligned with one another in such a way that a 3D-3D correspondence for all vertices is obtained (Blanz & Vetter, 1999). For example, if an index (e.g., k = 587) is assigned to the tip of the nose, then the nose tips for all the 3D face shapes will have the same index, k = 587. However, the x, y, and z coordinates that represent the nose-tip position on the faces tend to change. Owing to this fact, the number of p vertices corresponding to each face can be vectorized as

$$s_i = (x_{i1}, y_{i1}, z_{i1}, ..., x_{ip}, y_{ip}, z_{ip})^T , \quad (1)$$

where s_i has the dimension $n = 3 \times p$, n is the number of vertices, and $i = 1,...,m$ is the number of face shapes. The dimensions of the shape vectors are very large compared to the sample size, while the number of vertices, n, is equal to 75972, and the sample size m comprises 100 face shapes. If PCA is applied to the data, the covariance matrix will be $n \times n$, which is very huge. However, the same eigenvectors and eigenvalues can be derived from a smaller $m \times m$ matrix, while the covariance matrix C can be written as

$$C = \frac{1}{m} XX^T = \frac{1}{m} \sum_{i=1}^{m} x_i^T x_i \in R^{m \times m} , \quad (2)$$

where $X = [x_1, x_2, ..., x_m]^T \in R^{m \times n}$ is the data matrix with each vector, $x_i = s_i - s_0$ is centered around the mean s_0, and $i = 1,...,m$. PCA is then applied on the covariance matrix C. As a result of the analysis, a new shape vector, $s_{rec} \in R^n$, can be expressed as

$$s_{rec} = s_0 + E\alpha = s_0 + \sum_{i=1}^{m} \alpha_i e_i \ , \quad (3)$$

where $E=(e_1,e_2,...,e_m)$ is the matrix of scaled basis vectors, and α_i is the coefficient of the scaled basis vector e_i. Since E is an orthogonal matrix, the PCA coefficients α of the vector $x=s-s_0\in R^n$ can be derived from Eq. (3) as

$$\alpha = E^T x \ . \tag{4}$$

3D SHAPE MODEL FITTING TO NEW FACES

Learning models are trained from a set of examples to reach a state where the model will be able to predict the correct output for other examples. However, for this study, where the training data set is relatively small and there are too many missing features in the testing data, overfitting can easily occur. The goal of robust fitting algorithms is to reduce the chance of fitting noise and increase the accuracy in predicting new data. Noise in our case may occur due to intricacies in choosing the input feature points, which depend on the acquisition systems or the uncertainties imposed by the alignment methods. Fitting the shape model to a given 2D image is formulated as an optimization problem in solving the linear system in Eq. (3), while the goal of this inverse problem is to find the PCA-coefficients α rapidly and efficiently, given E and the shape vector x.

Given a number of feature points $f << p$, our problem is to find the 3D coordinates of all other vertices. In case of limited feature points, overfitting may occur by using approximation methods. In addition, using a holistic model such as a PCA-based model, the model cannot be adapted to the particular set of feature points; otherwise this would result in overfitting. Therefore, regularization can be used to assure the result is plausible according to the prior knowledge (Knothe, Romdhani & Vetter, 2006).

Assume that $s_f \in R^l (l=2f)$, s_{0f} shows the corresponding points on s_0 (the average 3D face shape) and that $x_f = s_f - s_{0f}$ is related to A such that

$$x_f = Ar + \varepsilon, \qquad A: R^m \mapsto R^l \,, \tag{5}$$

where $r \in R^m$ is the model parameter, and where $A \in R^{l\times m}$ is the corresponding subset of the matrix of scaled basis vectors $E^T \in R^{n\times m}$ and $\varepsilon \in R^l$ can be considered as measurement errors with unknown properties. Ultimately the goal is to estimate r as accurately as possible, given A and x_f. Since A is orthogonal, a simple least square technique can be used to solve the inverse problem as

$$r = (A^T A)^{-1} A^T x_f \,. \tag{6}$$

Because the measured feature points x_f capture only a small portion of the original image x with some measuring errors, regularization can be used as a constraint that utilizes the possible features in the holistic model to produce plausible results. The standard Tikhonov regularization enables an approximate solution which, can be written as

$$r_{reg} = (A^T A + \lambda I)^{-1} A^T x_f \,. \tag{7}$$

In our case, the original data matrix X is multivariate normal distribution whereas the means are zero, and the components are independent and have the same standard deviation. The authors also assume that the errors in x_f are independent with a zero mean and with the same standard deviation as in the original data. According to Bayes' theorem, under these assumptions the Tikhonov-regularized solution is the most probable solution (Vogel, 2002). However, the reconstructed face shapes have excessive smoothness. To improve the standard Tikhonov regularization, the stabilizing item is chosen to be the inverse of the diagonal eigenvalue matrix W to ensure that the solution will be within the boundary of the learning model. The model parameter α can be estimated as

$$\hat{\alpha} = (A^T A + \lambda W^{-1})^{-1} A^T x_f \,. \tag{8}$$

Then a new face shape, s_{rec}, can be obtained by applying $\hat{\alpha}$ to Eq. (3). Jiang et al. (2005) have used the same regularization equation in an iterative procedure in order to bring about a stable solution. In this work, the shape coefficients are calculated directly using Eq. (8).

TEXTURE WARPING

Before warping the 2D input face texture to the reconstructed 3D shape, the texture needs to be in the same coordinate system as the model texture is. Basically such registration is carried out with the guidance of feature points chosen from the given 2D input face images and corresponding feature points on the model face texture. To learn the deformation mapping between the input image and model textures in the database, 28 landmarks (including eyes, eyebrows, nose, mouth, and face contours) have been used. TPS is used to establish the mapping and interpolation for the deformation process. TPS is a commonly used basis function for representing coordinate mappings from rigid to nonrigid deformations. It is used for estimating a deformation function between two surfaces (Xiaoguang & Jain, 2008). Let g_0 and g_1 be two 2D shapes, and let $X = (x_1, x_2, ..., x_m) \subset g_0$ and $Y = (y_1, y_2, ..., y_m) \subset g_1$ be the correspondences (landmarks) between the two shapes, where m is the number of corresponding points. The warping function F that warps point set X into Y is given under the following condition

$$F(x_j) = y_j, \quad j = 1,2,...,m. \tag{9}$$

The interpolation deformation model is given in terms of the warping function F(x), where

$$F(x) = x.A + KW, \tag{10}$$

and where $x \in g_0$, A is an affine transformation, W is a fixed m-dimensional column vector of non-affine warping parameters constrained to $X^T W = 0$, and K is an m-dimensional row vector. K_i is the Green's function $\sigma(|x - x_i|)$, which can be written as $K_i(x) = (\sigma(|x - x_1|), \sigma(|x - x_2|), ..., \sigma(|x - x_m|))^T$. Some typical deformed 2D input faces registered with reference to model face texture using TPS are shown in FIGURE 1.

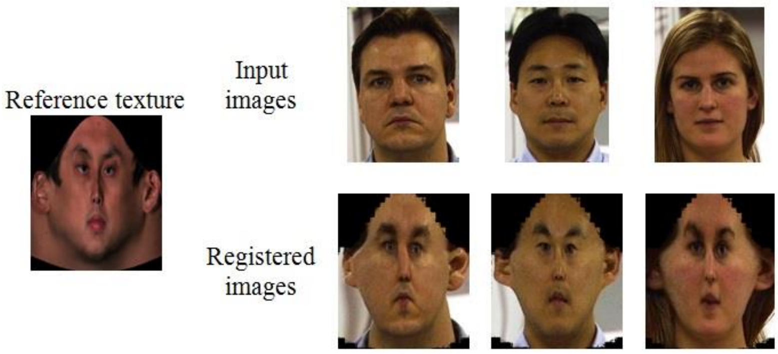

FIGURE 1. Typical 2D-2D registration based on TPS.

EXPERIMENTS AND DISCUSSION

This section intends to report the experimental evaluation aspects of the proposed techniques by visualizing some reconstructed results. The model is qualitatively evaluated pertaining to the visualization aspects of the reconstructed faces from their 2D face images. The USF Human ID 3D Face database (Blanz and Vetter, 1999), which contains 100 3D faces, has been used. The proposed model has been trained with the 100 3D face shapes. Each face shape has coordinates of 75972 vertices. They are aligned with one another as explained in (Levine & Yu, 2009). FIGURE 2 shows 3D face examples from the 3D database, including shape and texture.

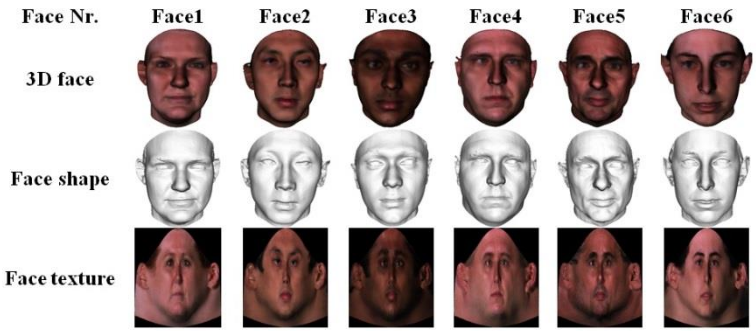

FIGURE 2. Examples of 3D faces from the USF Human ID 3D database.

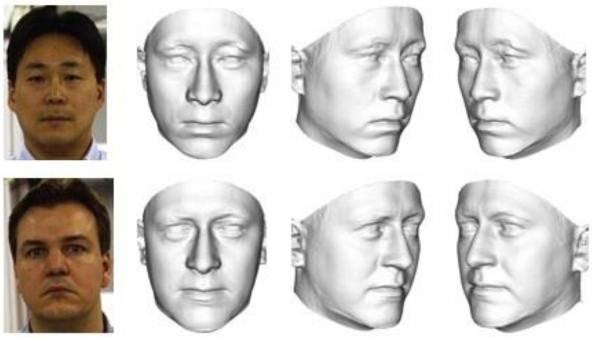
FIGURE 3. Examples of reconstructed 3D face shapes from their corresponding 2D faces using regularization.

RECONSTRUCTION OF 3D FACES SHAPES FROM IMAGES

The CMU-PIE database (Sim, Baker & Bsat., 2003) has been used for testing the visual effects of the proposed model. 3D models are intended to be reconstructed for the 2D images of the CMU-PIE database. From just a small number of 2D facial landmarks, the proposed algorithm can recover the 3D shapes of the given 2D face images. In this work, the input 2D images are in near-frontal pose with most of their expressions being neutral. The feature points which have been manually selected have been aligned with the reference 3D model using Procrustes Analysis; this is the usual preliminary step before the reconstruction stage. The aligned feature points have been used to compute the model parameters α using Eq. (8). Then α has been applied to Eq. (3) to reconstruct the 3D shape vector. FIGURE 3 shows two examples of reconstructed face shapes based on the proposed regularization algorithm.

WARPING THE TEXTURE ON THE RECONSTRUCTED SHAPES

28 landmarks (including eyes, eyebrows, nose, mouth, and face contours) have been used to register the input 2D image texture on the model texture. The resultant 2D textures (FIGURE 4, second column) are warped on the reconstructed 3D shapes. FIGURE 4 shows the texture-mapped results. It can be seen that, interestingly, the proposed model is capable of reconstructing 3D face shapes and do warping the original texture of the input image on the reconstructed 3D face shapes by retaining realistic facial features.

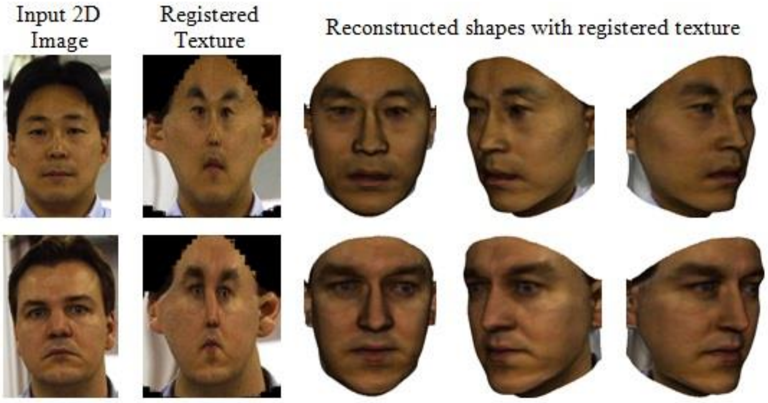

FIGURE 4. From original near frontal images, the regularized algorithm estimates 3D shapes from 78 feature points. The 2D input textures are first mapped on the model textures and then warped on the reconstructed shapes.

CONCLUSION

The problem of 3D facial modeling remains as a partially solved problem in the field of computer vision in terms of the accuracy and speed of reconstruction algorithms. This paper presented an efficient approach for recovering the 3D face of an individual given the 2D face image, in which prior knowledge has been subsumed. The standard Tikhonov regularization method has been extended by replacing the identity matrix with the eigenvalue matrix in order to solve the ill-posed problem of reconstructing complete 3D face shapes from 2D face images. The standard PCA model has been used to represent the object class of 3D face shapes for the reconstruction application. For color values of the input image, the proposed approach interpolates the input 2D textures with model texture and warps the interpolated textures to the reconstructed 3D face shapes. The proposed method has been tested using real 2D face images by visualizing the reconstructed results. For the texture warping, the 2D face deformation has been learned from the model texture using a set of facial landmarks. Thin Plate Spline (TPS) has been used for transferring the deformations based on those facial landmarks. 3D face shapes from the Human ID 3D database has been used to build the holistic PCA model. Some 2D face images from the CMU-PIE database have been used to visualize the reconstructed 3D faces. Our reconstructed results clearly demonstrate the effectiveness of the proposed method and retain the real characteristics of the given 2D face images. However, as the experiments have been done on near-frontal 2D face images, future work should focus on pose and expression images.

ACKNOWLEDGMENTS.

The first author would like to thank the Ministry of Higher Education, Malaysia, for supporting him with the MIS scholarship. This research is also partly supported by RU Grant 1001/PKCOMP/817055 (Universiti Sains Malaysia) and ERGS Grant 203/PKOMP/6730075 (Ministry of Higher Education, Malaysia).

REFERENCES

Blanz, V., Mehl, A., Vetter, T. and Seidel, H.P. 2004. A Statistical Method for Robust 3D Surface Reconstruction from Sparse Data. *Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission.* Thessaloniki, Greece: IEEE Computer Society, 293-300.

Blanz, V. and Vetter, T. 1999. A morphable model for the synthesis of 3D faces. *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. Anaheim, California: ACM Press/Addison-Wesley Publishing Co., 187-194.

Blanz, V. & Vetter, T. 2002. Reconstructing the Complete 3D Shape of Faces from Partial Information (Rekonstruktion der dreidimensionalen Form von Gesichtern aus partieller Information. *it+ti - Informationstechnik und Technische Informatik*, 44(6/2002):295-302.

Jiang, D., Hu, Y., Yan, S., Zhang, L., Zhang, H. and Gao, W. 2005. Efficient 3D reconstruction for face recognition. *Pattern Recognition,* 38(6)**:** 787-798.

Jing, L., Liu, S., Zhihong, L. and Meng, S. 2009. An image reconstruction algorithm based on the extended Tikhonov regularization method for electrical capacitance tomography. *Measurement*, 42(3):368-376.

Knothe, R., Romdhani, S. and Vetter, T. 2006. Combining PCA and LFA for surface reconstruction from a sparse set of control points. *Proceedings* of *the 7th International Conference on Automatic Face and Gesture Recognition*. Southampton, UK: IEEE Computer Society, 637-644.

Maghari, A., Liao, I. and Belaton, B. 2012. Quantitative analysis on PCA-based statistical 3D face shape modeling. *Computational Modelling of Objects Represented in Images III,* 13-18. CRC Press.

Sim, T., Baker, S. and Bsat, M. 2003. The CMU pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 25(12):1615-1618.

Vogel, C. R. 2002. *Computational methods for inverse problems*. Philadelpia: Society for Industrial and Applied Mathematics (SIAM).

Widanagamaachchi, W. N. and Dharmaratne, A. T. 2008. 3D Face Reconstruction from 2D Images. *Proceedings of the International Conference on Digital Image Computing: Techniques and Applications*. Canberra, Australia: IEEE Computer Society, 365-371.

Xiaoguang, L. and Jain, A. K. 2008. Deformation Modeling for Robust 3D Face Matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 30(8)**:** 1346-1357.

Ashraf Y. A. Maghari
Ibrahim Venkat
Bahari Belaton
School of Computer Sciences,
Universiti Sains Malaysia,
Penang, Malaysia,
myashraf2@gmail.com

Received: 11 August 2013
Accepted: 23 October 2013
Published: 1 March 2014