Proceedings of the International MultiConference of Engineers and Computer Scientists 2011 Vol I,
IMECS 2011, March 16 - 18, 2011, Hong Kong

The Novel Approach for 3D Face Recognition Using Simple Preprocessing Method

Parvin Aminnejad[1] , Ahmad Ayatollahi[2] , Siamak Aminnejad[3], Reihaneh Asghari

Abstract— **In this work, we presented a novel approach for automated 3D face recognition using range data. The contributions of the paper can be summarized as: 1) data registration, 2) data comparison. In first step, the nose tip was used as the reference point and three dimensional face shape was normalized to standard image size. Then 2DPCA was applied to the resultant normalized data and the output data were used as the feature vectors. The Support Vector Machine was used in classification step. Recognition rate of 98% was achieved.**

Index Terms—**3D face recognition, 2DPCA, SVM**

I. INTRODUCTION

Automatic recognition of human faces from 2D intensity images has been studied extensively in the computer vision community. Although significant progress has been made, the task of automated, robust face recognition is still a distant goal. Two-dimensional image-based methods are inherently limited by variability in imaging factors such as illumination and pose. These problems of 2D face recognition systems are solved by using the 3D face shapes, since 3D data is not affected by translation and rotation and is immune to the effect of illumination variation. The 3D face recognition research is, however, still weakly reported in the published literature. Cartoux et al. [4] approach 3D face recognition by segmenting a range image based on principal curvature and finding a plane of bilateral symmetry through the face. This plane is used to normalize for pose. They consider methods of matching the profile from the plane of symmetry and ofmatching the face surface. Gordon [3] also begins with a segmentation based on mean and Gaussian curvature. The nose region and ridge and valley lines from the segmentation are used to register the image to a standard pose. Matching is then done by computing the volume difference between registered probe and gallery surfaces. Tanaka et al. [1] also perform curvature-based segmentation and represent the face using an Extended Gaussian Image (EGI). Recognition is then performed using a spherical correlation of the EGIs. Moreno and co-workers [5] approach 3D face recognition by first performing a segmentation based on Gaussian curvature and then creating a feature vector based on the segmented regions. They report results on a dataset of 420 face meshes representing 60 different persons, with some sampling of different expressions and poses for each person. They report 78% success in recognition on the subset of frontal views. One limitation to some existing approaches to 3D face recognition involves sensitivity to size variation. Approaches that use a purely curvature-based representation, such as extended Gaussian images, are not able to distinguish between two faces of similar shape but different size. Approaches which use a PCA type algorithm can handle size change between faces. An object recognition system generally consists of two main parts: data registration and data comparison. The accuracy of registration will greatly impact on the result of following comparison. Although Blanz [6] gave a nice solution to registration of 3D facial data, but it’s high time-cost made it hard to incorporate into a practical recognition system. In this paper, we present a size and expression invariant recognition system based on the two-dimensional principal component analysis. In order to discard the background information, we apply a threshold on Z coordinate values of a 3D facial image. Also we propose a simple and fast registration method; the detected face shape is registered coarsely and then, in the fine registration step, it is normalized to a standard image of size pixels in such a way that the nose tip point is selected to be the image center. Our proposed method works well with a high resolution 3D facial data (nearly 18000 points). Our experiments were performed on 3D CASIA database [2]. CASIA database contains 4674 three-dimensional facial surface images corresponding to 123 individuals, and there are 38 different images per each person.

*P.Aminnejad, is with faculty of Electrical, Computer and IT Engineering Islamic AZAD University –Qazvin Branch ,Qazvin, Iran (phone:0098-21-22595534,email:parvin_amin2003@yahoo.com).
** A.Ayatollahi is with the Electrical Engineering Department , Iran ,University of Scince and Tecnology (IUST). Tehran,(Tel:0098-21-77451502,email:ayatollahi@iust.ac.ir).

II. OVERVIEW OF OUR PROPOSED RECOGNITION SYSTEM

In this work, we propose a new technique for human face recognition problem in 3D images with the ability of handling different expressions of one's facial image. A sample 3D face image from CASIA database is depicted in Fig. 1. Also an overview of our proposed algorithm is depicted in Fig. 2.

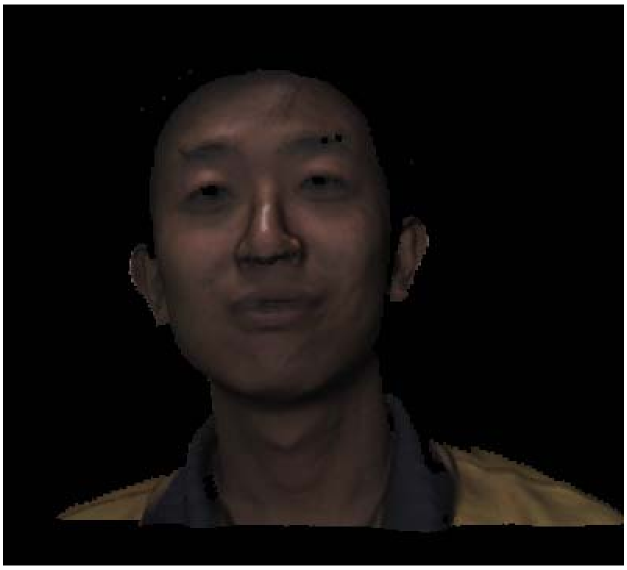

Figure 1. A sample 3D face image from CASIA database

III. FACE DETECTION

A facial image is first subjected to a face detection stage in which the background and some unnecessary regions such as sides of the head, neck, and ears are eliminated by thresholding the Z coordinate values of the 3D rang data. For this purpose we use Otsu's Method [7]. Otsu suggested a criterion by which the best threshold for images with bimodal histogram can be determined. The criterion states that the threshold should be chosen in such a way that minimizes the weighted sum of within group variances for the two groups that result from separating the gray levels at the threshold value. Fig. 3 illustrates the 3D mesh representation of face image from CASIA database and the result of face detection depicted in Fig. 4.

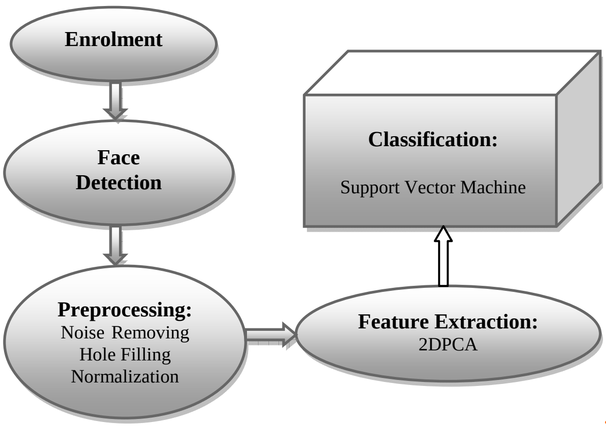

Figure 2. Proposed system Overview

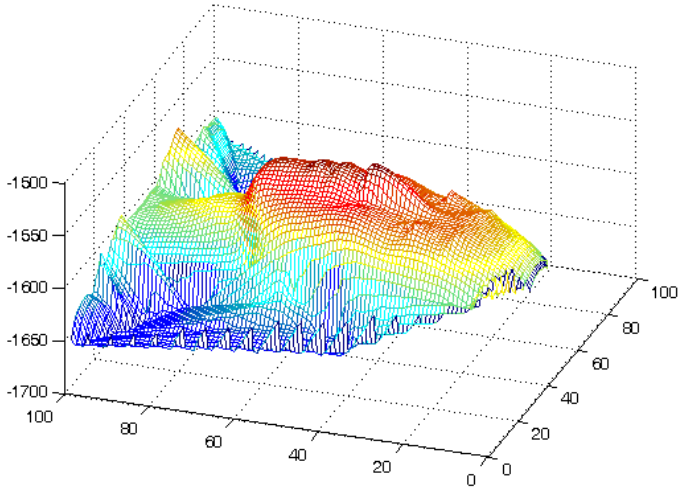

Figure 3. 3D mesh representation of face image from CASIA database

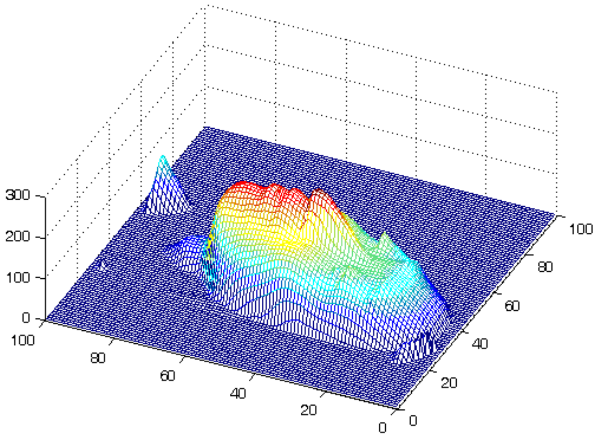

Figure 4. Result of face detection

IV. PREPROCESSING

Preprocessing is an important stage of the recognition systems, since all the features will be extracted from the output of this step

A. Noise reduction & hole filling

We used the simple technique in the preprocessing step: removing spike noise and hole filling. We used median filtering for cancelling of spike noise and 2D interpolation using of cubic method for hole filling purpose.

B. Normalization

This step will permit to obtain the three normalized coordinates for each point. So we need a reference point. The best reference point is the end and the top of the nose of frontal image. In most images, the nose is the closest part of the face to the 3D scanner, so it has the highest depth value among all the facial points. By using a window that calculates the sum of the depth values of its corresponding pixels, the nose is detected as the coordinates of the central pixel of the window with the maximum value. After

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

detecting the nose, all images in the database are normalized to a standard image and then aligned so that the nose lies exactly in the center of each image at the (50, 50) x-y coordinate. The result of this step depicted in Fig. 5

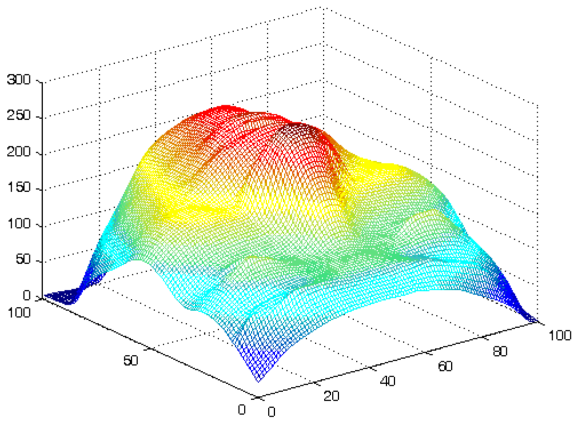

Figure 5. Result of Normalization

V. FEATURE EXTRACTION

The feature extraction stage employs 2DPCA on the normalized range image and the resultant eigen-images are used as the feature vectors for the matching process. 2DPCA is equivalent to a special case of an existing feature extraction method, block-based PCA, which has been used for face recognition in a number of systems. It was developed for image feature extraction based on 2D matrices as opposed to the standard PCA which is based on 1D vectors. An inter-image covariance matrix is constructed using the 2D matrix representation of each image, which results in an ensample covariance matrix of the training set. Let the training set of 3D facial image be $F_1 F_2, \dots, F_M$ each of which is of size $n \times n$. The average face of the set is defined as:

$$\bar{F} = \frac{1}{M}\sum_{i=1}^{M} F_i \qquad (4)$$

The image covariance matrix is given by:

$$Cov = \frac{1}{M}\sum_{i=1}^{M}(F_i - F)^T (F_i - F) \qquad (5)$$

The optimal number (m) of projection set of an image F, denoted by X, is defined through $y_k = F.X_k \ (k = 1, \dots m)$ [9].

To achieve the maximum scatter of the projected feature vectors the projection axes $(X_1 X_2, \dots, X_m)$ have to be the orthonormal eigenvectors of the covariance matrix Cov corresponding to the first m largest eigenvalues $(Y_1 Y_2, \dots, Y_m)$. The projected vectors are used as feature vectors in our implementation. By reconstructing some sample images from the database, the value of m=12 was determined experimentally in which the recognition rate is maximize.

VI. CLASSIFICATION

We use SVM for classification purpose. For details of SVM see Vapnik [10,11]. SVM is a binary classification method that finds the optimal linear decision surface based on the concept of structural risk minimization. The decision surface is a weighted combination of elements of the training set. These elements are called support vectors and characterize the boundary between the two classes. The input to a SVM algorithm is a set $\{(x_i, y_i)\}$ of labeled training data, where is the data and $y_i = +1$ or $y_i = -1$ is the label. The output of a SVM algorithm is a set of N_s support vectors s_i, coefficient weights α_i, class labels y_i of the support vectors, and a constant term b. The linear decision surface is:

$$W.Z + b = 0 \qquad (6)$$

Where

$$W = \sum_{i=1}^{N_s} \alpha_i y_i s_i \qquad (7)$$

SVM can be extended to nonlinear decision surfaces by using a kernel $k(.,.)$ that satisfies Mercer's condition [10, 11]. The nonlinear decision surface is

$$\sum_{i=1}^{N_s} \alpha_i y_i k(s_i, z) + b = 0 \qquad (8)$$

A facial image is represented as a vector $Y \in \mathcal{R}^N$ where $\mathcal{R}^N$ is referred to as face space. We obtain the face space with projecting the facial image on the eigenvectors generated by performing 2DPCA on a training set of faces. In identification, there is a gallery $\{g_i\}$ of m known individuals. The algorithm is presented with a probe p to be identified. The first step of the identification algorithm computes a similarity score between the probe and each of the gallery images. The similarity score between p and g_i is

$$\delta_j = \sum_{i=1}^{N_s} \alpha_i y_i k(s_i, g_j \quad p) + b \qquad (9)$$

In the second step, the probe is identified as person that has minimum similarity score δ_j .

VII.EXPERMENTAL RESULT

As mentioned in section 5, we are used the projected vectors $(Y_1, Y_2, \ldots, Y_m)$ as feature vectors in our implementation. Figure 7 shows a plot of the magnitude of the eigenvalues versus their rank. The magnitude of an eigenvalue is equal to the variance in the data set that is spanned by its corresponding eigenvector. Thus, it is obvious that higher-order eigenvectors account for less energy in the approximation of the data set since their eigenvalues have low magnitudes. We demonstrate SVM-based identification algorithm on 4674 frontal images from the CASIA database of 3D facial images. In the implementation, a radial basis kernel was used.

CASIA database contains 4674 three-dimensional facial surface images corresponding to 123 individuals, and there are 38 different images per each person. We were divided these 3D faces into disjoint training and testing sets. Each training and testing sets consisted of nine and seven images of 123 people respectively. Figure 8 shows the percentage of recognition vs. the number of eigenvectors.

Table 1 summarizes the recognition rate for a four training samples employed in experimentation and Figure 9 shows the plot of the data in that table. Since we have applied smoothing stage to images, the curvature of images with expression appears natural, so our system is expression invariant.

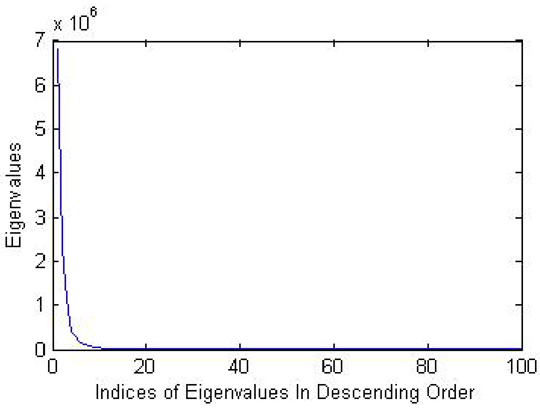

Figure 6. Magnitude of the eigenvalues versus their rank.

Table 1.Recognition rate of the system at different number of training images

Number of training images	9	10	11	12
Recognition rate	0.91	0.94	0.96	0.98

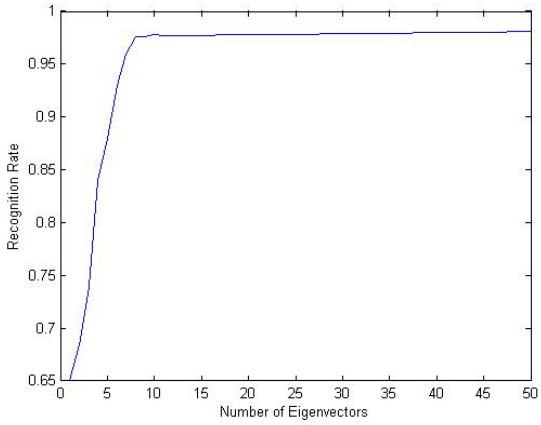

Figure 7. Percentage of recognition vs. the number of eigenvectors.

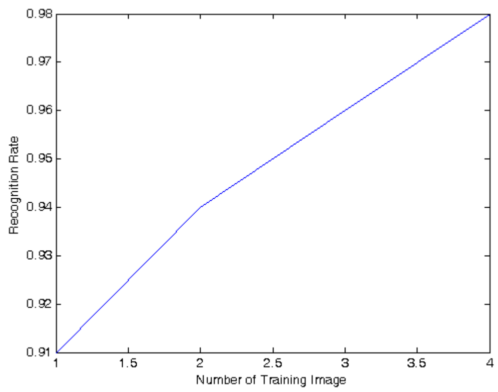

Figure 8. A plot of the recognition rate vs. the number of training images.

VIII. CONCLUSION

We presented a novel approach for automated 3D face recognition using range data. A facial image is first subjected to a face detection stage in which the background and some unnecessary regions such as sides of the head, neck, and ears are eliminated by thresholding the Z coordinate values of the 3D rang data. In order to perform a scale-invariant identification, facial depth-values were scaled between 0 and 255. The resultant facial image map was

normalized and then aligned with the coordinate system centered at the subject nose. It was then smoothed to achieve an expression invariant recognition system.

The 2DPCA is used for feature extraction. The proposed method was tested on the CASIA database.

Classification was carried out by calculating the similarity score between the feature vectors. The SVM classifier was used in choosing the closest match. Recognition rate of 98% was achieved.

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

REFERENCES

[1] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical ...Third International Conference on Automated Face and Gesture Recognition, pages 372–377, 1998.

[2] CASIA DATABASE

[3] G. Gordon. Face recognition based on depth maps and surface curvature. Geometric Methods in Computer Vision, SPIE 1570:1–12, July 1991.

[4] J. Y. Cartoux, J. T. LaPreste, and M. Richetin. Face authentication or recognition by profile extraction from range images. Proc. of the Workshop on Interp. of 3D Scenes, pages 194–199, November 1989.

[5] A. B. Moreno, Angel Sanchez, J. F. Velez, and F. J. Diaz. Face recognition using 3D surface-extracted descriptors. Irish Machine Vision and Image Processing Conference (IMVIP2003), September 2003.

[6] Volker Blanz and Thomas Vetter, “A morphablemodel for the synthesis of 3d faces,” in Computer Graphics Pmceedings SIGGRAPH‘99,1999, pp. 187-194.

[7] N. Otsu, "A threshold selection method from gray-level histogram," IEEE Transactions on Systems, Man and Cybernetics, Vol.8, No.1, pp.62-66, 1979

[8] Tasdizen, T., Whitaker, R., Burchard, P., and Osher, S. 2002. Geometric surface smoothing via anisotropic diffusion of normals. In Proc. IEEE Conf. Visualization, pp. 125–132.

[9] Yang, D. Zhang, A. F. Frangi, and J. Y. Yang, "Twodimensional PCA: A new approach to appearance-based face representation and recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vo1.26, No.1, pp.131-137, January 2004.

[10] V. Vapnik. The nature of statistical learning theory. Springer, New York, 1995.

[11] C. J. C. Burges. A tuturial on support vector machines for pattern recognition. Data mining and knowledge discovery, (submitted), 1998.

ISBN: 978-988-18210-3-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)