ISSN No. 2278-3091
Volume 2, No.2, March - April 2013
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://warse.org/pdfs/2013/ijatcse06222013.pdf

3D Face Recognition based on Deformation Invariant Image using Symbolic LDA

P. S. Hiremath[1], Manjunatha Hiremath[1]
[1]Department of Computer Science
Gulbarga University, Gulbarga-585106
Karnataka, India.

ABSTRACT

Face recognition is one of the most important abilities that the humans possess. There are several reasons for the growing interest in automated face recognition, including rising concerns for public security, the need for identity verification for physical and logical access to shared resources, and the need for face analysis and modeling techniques in multimedia data management and digital entertainment. In recent years, significant progress has been made in this area, with a number of face recognition and modeling systems have been developed and deployed. However, accurate and robust face recognition still offers a number of challenges to computer vision and pattern recognition researchers, especially under unconstrained environments. In this paper, a novel deformation invariant image based 3D face recognition is proposed. The experiments are done using the 3D CASIA Face Database, which includes 123 individuals with complex expressions. Experimental results show that the proposed method substantially improves the recognition performance under various facial expressions.

Key words : 3D face recognition, facial expression, geodesic distance, symbolic LDA, deformation invariant image.

1. INTRODUCTION

Face recognition based on 3D surface matching is a promising method for overcoming some of the limitations of current 2D image-based face recognition systems. The 3D shape is generally invariant to the pose and lighting changes, but not invariant to the non-rigid facial movement, such as expressions. Collecting and storing multiple templates to account for various expressions for each subject in a large database is not practical. Current 2D face recognition systems can achieve good performance in constrained environments. However, these systems still encounter difficulties in handling large amounts of facial variations due to head pose, lighting conditions and facial expressions. Since the 2D projection (image or appearance) of a three-dimensional human face is sensitive to the above changes, the approach that utilizes 3D facial information appears to be a promising avenue to improve the face recognition accuracy. According to the evaluation of commercially available and mature prototype face recognition systems provided by face recognition vendor tests (FRVT) [1,2], the recognition performance under the unconstrained condition is not satisfactory. In this paper, we at-tempt to realize 3D face recognition robust to expression variations.

In fact, the human face contains not only 2D texture information but also 3D shape information. Recognition using 2D images results in the loss of some information. An alternative idea is to represent the face or head as a realistic 3D model, which contains not only texture and shape information, but also structural information for simulating facial expressions. In addition, some techniques in computer graphics can be explored to simulate facial variations, such as expressions, ageing and hair style, which provide an ideal way to identify a changing individual.

With the rapid development of 3D acquisition equipment, face recognition based on 3D information is attracting more and more attention [3,6]. In 3D face recognition, depth information and surface features are used to characterize an individual. This is a promising way to understand human facial features in 3D space and to improve the performance of current face recognition systems. Moreover, some current 3D sensors can simultaneously obtain texture and depth information, resulting in multi-modal recognition [4,5]. This makes 3D face recognition a promising solution to overcome difficulties in 2D face recognition. Variations in illumination, expression and pose are the main factors influencing face recognition performance. For 3D face recognition, illumination variations do not influence the recognition performance that much. This is not strange since the original facial data are usually captured by a laser scanner which is robust to illumination variations. Pose variations only affect the recognition performance a little bit because some registration methods can accurately align the face data, thus reducing the influence of pose variations. Expression variations greatly affect the recognition performance. Expression variations distort the facial surface, and result in the change of the facial texture. Moreover, expression variations also cause the registration error to increase. It is noted that the expression variation is one of the most difficult factors in 3D face recognition.

There also exist some attempts to overcome the expression variation in 3D face recognition field. Facial surface varies differently during expression changes: some regions are deformed largely and others little. In [5,7] divided the whole facial surface into some sub-regions. The rigid regions around the nose area were used to be matched and combined to perform the recognition. But it was hard to efficiently determine the rigid regions across expression changes. Making deformable 3D face model is another way to simulate the facial expression. In [8] authors extracted the deformation from a certain controlled group of face data. Then, the extracted deformation was transferred to and synthesized for the neutral models in the gallery, thus yielding deformed templates with expressions. The comparison between the tested face data and the deformed templates was performed to finish the recognition. In [9,10], authors used an annotated face model to fit the changed facial surface, and then obtained the deformation image by the fitted model. A multi-stage alignment algorithm and the advanced wavelet analysis resulted in robust performance. In these studies, the main problem has been how to build a parameterizing 3D model from optical or range images, which is not well solved. Facial expression deforms the facial surface in a certain way, which can be used in face recognition. In [11], authors represented a facial surface based on geometric invariants to isometric deformations and realized multi-modal recognition by integrating flattened textures and canonical images, which was robust to some expression variations. In [12], authors proposed a geodesic polar representation of the facial surface. This representation tried to describe the invariant properties for face recognition under isometric deformations of the facial surface. Face matching was performed with surface attributes defined on the geodesic plane. Also based on the assumption of isometric transformation, [13] proposed a novel representation called ‘‘isoradius contours” for 3D face registration and matching. Here in, an isoradius contour was extracted on the 3D facial surface that was a known fixed Euclidean distance relative to certain predefined reference point (the tip of the nose). Based on these contours, a promising result of 3D face registration could be achieved. Empirical observations show that facial expressions can be considered as isometric transformations, which do not stretch the surface and preserve the surface metric.

In this paper, the objective is to propose a new representation of deformation invariant image that uses the radial geodesic distance to realize the face recognition, which is robust to expressions. First, the depth image and the intensity image from the 3D face database are obtained. Then, geodesic level curves are generated by constructing radial geodesic distance image from the depth image. Finally, deformation invariant image is constructed by evenly sampling points from the selected geodesic level curves in the intensity image. Further, Symbolic LDA method is used for classification in the face recognition system.

2. MATERIALS AND METHODS

For experimentation, we consider large 3D face database, namely, 3D CASIA face database [14], which is used to test the proposed algorithm. The 3D images were collected indoors during August and September 2004 using a non-contact 3D digitizer, Minolta VIVID 910. This database contains 123 subjects, and each subject has not only separate variation of expressions, poses and illumination, but also combined variations of expressions under different lighting and poses under different expressions. Some examples with expression variations are shown in Figure . 4. These variations provide a platform on which the performance of 3D face recognition will be investigated under different variations. We use 1353 images from this database (11 images for each person) in our experiments. These images are divided into three subsets, that is, the training set, the gallery set and the probe set. The training set contains 253 images, corresponding to the last 23 of the 123 subjects, 11 images for each person. The gallery set contains 100 images from the first image of the other 100 subjects (under the condition of front view, office lighting, and neutral expression). The other 1000 images from the above 100 subjects are used as the probe set [14].

The probe set is further divided into four subsets:

- EV probe set (200 images): the probe set including closed-mouth expression variations, such as anger and eye closed.
- EVO probe set (300 images): the probe set including opened-mouth expression variations, such as smile, laugh, and surprise.
- EVI probe set (200 images): the probe set including closed-mouth expression variations under side lighting, such as anger and eye closed.
- EVIO probe set (300 images): the probe set including opened-mouth expression variations under side lighting, such as smile, laugh, and surprise.

The EV and EVI probe sets include the facial expression with closed mouth, and the EVO and EVIO probe sets include the expression with opened mouth. They have different recognition performances, which will be demonstrated by the experimental results.

3. PROPOSED METHODOLOGY

The proposed method consists of preprocessing, building, deformation to invariant image, based on radial geodesic distance for extracting the robust features across expression changes and Symbolic LDA for face recognition. It combines the shape and the texture information in the 3D face effectively.

3.1 Preprocessing

The 3D data in our study consist of a range image with texture information as shown in Figure .4. In this section, we preprocess these original 3D data to obtain the normalized depth and intensity images. First, we detect the nose tip in the range image. Different from 2D color or intensity images, the nose is the most distinct feature in the facial range image. The method proposed by [18] is used to localize the nose tip. This algorithm utilizes two local surface properties, i.e. local surface feature and local statistical feature. It is fully automated, robust to noisy and incomplete input data, immune to rotation and translation and suitable to different resolutions. Alignment is performed according to the method by [19]. A front facial 3D image with neutral expression is selected as the fixed model, and all the other 3D images are rotated and translated to align with it. Based on the detected nose tip, all the range images are translated and coarsely aligned together. Then, the classic algorithm of the iterative closest point (ICP) [15] is used to further register them. The 3D data that are being processed have the same size as the real subjects. We use a 160x128 pixels rectangle centered at the detected nose tip to crop the original 3D data.

3.2 Expression Invariant Image:

In the 3D facial data, the tip of the nose can be detected reliably. Therefore, we regard this point as the reference point, and calculate the geodesic distance from it to other points. Since all the points are centered around the nose tip and the geodesic distance is in the radial direction, it is called as radial geodesic distance. The computation of this kind of distance is described as following: The surface curve between two given points can be described as

$$l(t) = (x(t), y(t), z(t))$$

where $x(t)$ and $y(t)$ refer to the position in X-Y plane and $z(t)$ refers to the corresponding depth value. The geodesic length *d* of this curve is given by:

$$d = \int_a^b \sqrt{x_t^2 + y_t^2 + z_t^2 dt}$$

where the subscripts denote partial derivatives, e.g. $x_t = {}^{dx}\!/_{dt}$. The radial geodesic distance can be approximately computed with sum of the lengths of the line segments embedded in the facial surface.

The Deformation invariant image based on the radial geodesic distance is described as follows: The nose tip has been determined in the normalized depth image. The radial geodesic distance between the nose tip and any other pixel is computed. The computation program evolves along an emissive shape as shown in Figure 1. Thus, one geodesic image is obtained as shown in Figure 2(a). In Figure 2(b), darker intensities mean smaller distance. Further, we can obtain the geodesic level curves are obtained, each of which consists of pixels with the same radial geodesic distance to the nose tip. Two neighboring level curves may have the identical change of geodesic distance or the log distance. In the present study, the geodesic distance is used. Figure 2(c) shows the geodesic level curves. Since the elliptical mask is used, some level curves end along the elliptical edge. We apply the level curves to the intensity image which corresponds to the depth image as shown in Figure . 2(d). The level curves determine the sampling position in the intensity image. The sampled pixels form a new image representation, which is called deformation invariant image. It is noted that different depth images of the same person have different geodesic level curves due to expression variations. Different level curves determine different sampling positions in intensity images. With the assumption of isometric deformation, these sampling positions balance the deformation of expressions in intensity images.

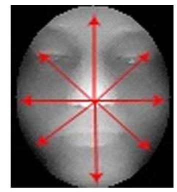

Figure . 1. Emissive shape for computing geodesic distance.

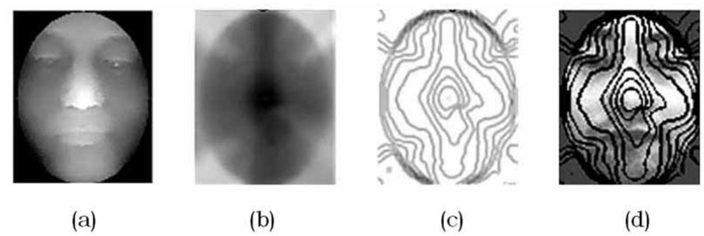

Figure . 2. Deformation invariant image. (a) Depth image; (b) geodesic distance image; (c) Geodesic level curves; (d) sampling position in the intensity image using geodesic level curves.

3.3 Symbolic LDA

We consider the extension of linear discriminant analysis (LDA) to symbolic data analysis frame work [12,14,15]. Consider the 3D range face images $\Gamma_1, \Gamma_2, ..., \Gamma_n$, each of size M x N, from 3D range face image database. Let $\Omega = \Gamma_1, \Gamma_2, ..., \Gamma_n$ be the collection of n 3D range face images of the database, which are first order objects. Each object $\Gamma_l \in \Omega,\ l = 1, 2, ..., n$, is described by a matrix A_l $(l = 1, 2, ..., n)$, where each component $\tilde{Y}_{ab}, a = 1, 2, ..., M$, and b=1,2,…,N, is a single valued variable representing the 3D range values of the face image Γ_l. An image set is a collection of face images of m different subjects and each subject has different images with varying expressions and illuminations. Thus, there are m number of second order objects (face classes) denoted by $E = \{c_1, c_2, ..., c_m\}$, each consisting of different individual images, $\Gamma_l \in \Omega$, of a subject. The face images of each face class are arranged from right side view to left side view. The feature matrix of k^{th} sub face class c_i^k of i^{th} face class c_i, where $k = 1, 2, ..., q$, $i = 1, 2, ..., m$, is described by a matrix X_i^k of size $M \times N$ that contains interval variable a_{iab}^k, $a = 1,2,...,M$, and $b = 1,2,...,N$. The matrix is called as symbolic face and is represented as :

$$X_i^k = \begin{bmatrix} a_{i11}^k & \dots & a_{i1N}^k \\ . & . & . \\ a_{iM1}^k & \dots. & a_{iMN}^k \end{bmatrix}$$

The interval variable a_{iab}^k of k^{th} sub face class c_i^k of i^{th} face class c_i is described as $a_{iab}^k(c_i^k) = [\underline{x}_{iab}^k, \overline{x}_{iab}^k]$, where $\underline{x}_{iab}^k$ and $\overline{x}_{iab}^k$ are minimum and maximum intensity values, respectively, among $(a,b)^{th}$ feature inside the k^{th} sub face class of i^{th} face class. Thus, we obtain the qm symbolic faces from the given image database[16,17].

Now, we apply LDA method to the centers $x_{iab}^k \in \mathbb{R}$ of the interval $[\underline{x}_{iab}^k, \overline{x}_{iab}^k]$ given by

$$x_{iab}^{k^c} = \frac{\overline{x}_{iab}^k + \underline{x}_{iab}^k}{2}$$

The $M \times N$ symbolic face $X_i^{k^c}$ containing the centers $x_{iab}^{k^c} \in R$ of the intervals a_{iab}^k of symbolic face X_i^k is given by

$$X_i^{k^c} = \begin{bmatrix} a_{i11}^{k^c} & \dots & a_{i1N}^{k^c} \\ . & . & . \\ a_{iM1}^{k^c} & \dots. & a_{iMN}^{k^c} \end{bmatrix}$$

In the symbolic LDA approach, to calculate the scatter (within and between class) matrices of qm symbolic faces X_i^k, where i=1,2,…,m and k=1,2,…,q, we define the within-class image scatter matrix S_w as

$$S_w = \sum_{i=1}^{m} \sum_{k=1}^{q} (X_i^{k^c} - M_i)^T (X_i^{k^c} - M_i)$$

where $M_i = \frac{1}{q} \sum_{k=1}^{q} X_i^{k^c}$, and the between-class image scatter matrix S_b as

$$S_b = \sum_{i=1}^{m} (M_i - M)^T (M_i - M),$$

where $M = \frac{1}{qm} \sum_{i,k} X_i^{k^c}$. In discriminant analysis, we want to determine the projection axis that maximizes the ratio $\frac{\det\{S_b\}}{\det\{S_w\}}$. In other words, we want to maximize the between-class image scatter matrix while minimizing the within-class image scatter matrix. It has been proved that this ratio is maximized when the column vector of projection axis V is the eigenvector of $S_w^{-1} S_b$ corresponding to first *p* largest eigenvalues. For each symbolic face X_i^k, the family of projected feature vectors, $Z_1, Z_2, ..., Z_p$ are considered as:

$$Z_s = X_i^k V_s$$

where s=1,2,…,p. Let $B_i^k = [Z_1, Z_2, ..., Z_p]$, which is called as the feature matrix of the symbolic face X_i^k. The feature matrix B_{test} of the test image X_{test} is obtained as :

$$Z_{(test)s} = X_{test} V_s \ ,$$

where s=1,2,…,p and $B_{test} = [Z_{(test)1}, Z_{(test)2}, ..., Z_{(test)p}]$.

3.4 Proposed Method

In the recognition system, the image is usually presented by one vector. Here, the deformation invariant image is also converted into one vector. Different vectors from different images should have the same dimensionality and corresponding components. To meet these requirements, the sampling rule is made for all images. First, geodesic distance image is obtained from the depth image, and then geodesic level curves are computed. The intensity image is further sampled by using the level curves, and the deformation invariant image is then constructed. Finally, the deformation invariant image is converted into one vector for recognition. In fact, the same position in different deformation invariant images consists of intensity pixels, which have the same radial geodesic distance to the nose tip. This proposed representation is invariant to the facial surface deformation and is expected to be robust to expression variations. The Figure 3 shows the overview of proposed framework.

4. RESULTS AND DISCUSSION

For experimentation, we consider the 3D CASIA face database. The proposed method is implemented using Intel Core 2 Quad processor @ 2.66 GHz machine and MATLAB 7.9. In the training phase, 11 frontal face images, with different expressions, of each of the 123 subjects are selected as training data set. For each face class (subject), two

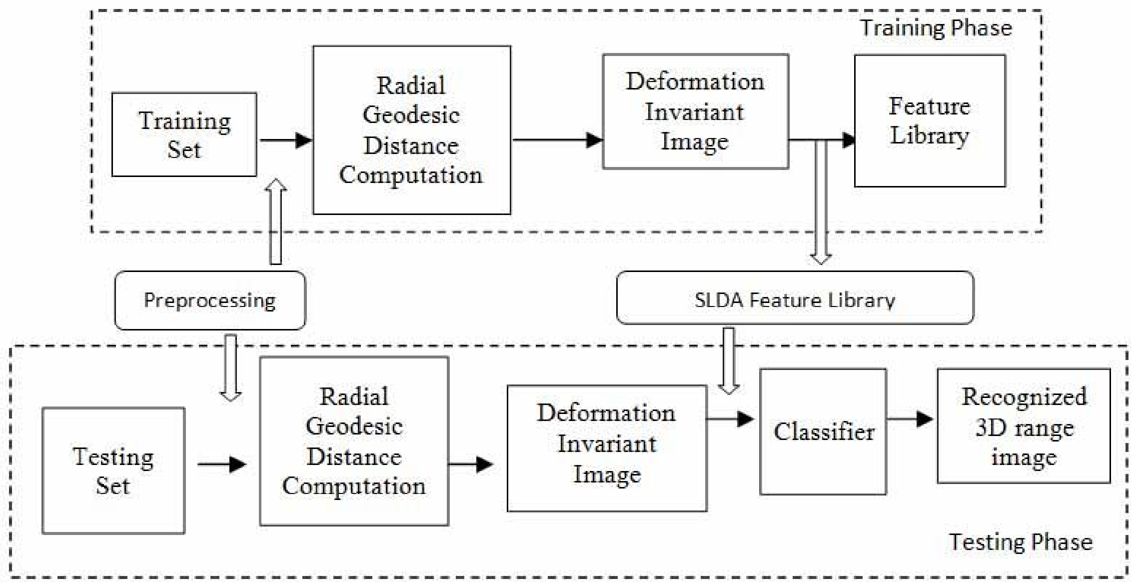

Fig. 3. Overview of proposed framework

Fig 4. Expression variations of the CASIA 3D Face Database

subclasses are formed; one subclass contains the face images with varying illumination, while the other subclass contains the face images of the same subject with varying facial expressions. In the testing phase, randomly chosen 200 face images of the 3D CASIA face database with variations in facial expressions are used. The sample training images which are used for experimentation are shown in the Figure . 4. The reocgnition performance in terms of accuracy and time is given in the table 1, which compares well with the methods in the literature. The recognition accuracy of 99.60% is achieved by the proposed method.

TABLE I. RANK-ONE RECOGNITION ACCURACY(%) FOR COMPARISON IN DIFFERENT PROBE SETS.

Probe Sets	Proposed Method		[20] Li2008	[11] Bron2007
	Recognition accuracy	Time Taken (In Secs.)		
EV	99.60%	2.953	94.5%	95.0%
EVO	91.70%	2.948	90.3%	89.7%
EVI	88.60%	2.935	88.0%	87.5%
EVIO	87.10%	2.930	85.3%	86.7%

5. CONCLUSION

In this paper, we have proposed a novel method for three dimensional (3D) face recognition using Radial Geodesic Distance and Symbolic LDA based features of 3D range face images. In this method, the Symbolic LDA based feature computation takes into account face image variations to a larger extent and has advantage of dimensionality reduction. The experimental results have yielded 99.60% recognition performance with reduced complexity and a small number of features, which compares well with other state-of-the-art methods. The experimental results demonstrate the efficacy and the robustness of the method to facial expression variations. The recognition accuracy can be further improved by considering a larger training set and a better classifier.

ACKNOWLEDGEMENT

The authors are indebted to the University Grants Commission, New Delhi, for the financial support for this research work under UGC-MRP F.No.39-124/2010 (SR).

REFERENCES

[1] Zhao. W., Chellappa R., Philips P. J., Rosenfeld A., “Face Recognition : A literature survey”, ACM Comput. Surveys (CSUR) Archive 35 (4), 399-458.

[2] Phillips, P.J., Scruggs, W.T., O’Toole, A.J., Flynn, P.J., Bowyer, K.W., Schott, C.L., Sharpe, M., Report of FRVT 2006 and ICE 2006 Large-Scale Results, Tech. Rep. NISTIR 7408. (2007)

[3] Bowyer, K., Chang, K., Flynn, P., 2006. A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition. Comp. Vis. Image Understand. 101 (1), 1–15.

[4] Wang, Y., Chua, C., Ho, Y., Facial feature detection and face recognition from 2D and 3D images. Pattern Recognition Lett. 23, 1191–1202. (2002)

[5] Chang, K.I., Bowyer, K.W., Flynn, P.J., 2005a. An evaluation of multi-modal 2D + 3D biometrics. IEEE Trans. PAMI 27 (4), 619–624. (2005)

[6] Tsalakanidou, F., Malassiotis, S., Strintzis, M.G., 2005. Face localization and authentication using color and depth images. IEEE Trans. Image Process. 14 (2), 152–168.

[7] Chang, K.I., Bowyer, K.W., Flynn, P.J., 2006. Multiple nose region matching for 3D face recognition under varying facial expression. IEEE Trans. PAMI 28 (10), 1695–1700.

[8] Lu, X., Jain, A.K., 2006. Deformation modeling for robust 3D face matching. In: Proc. CVPR'06, pp. 1377–1383.

[9] Passalis, G., Kakadiaris, I.A., Theoharis, T., Toderici, G., Murtuza, N., 2005. Evaluation of 3D face recognition in the presence of facial expressions: An annotated deformable model approach. In: Proc. FRGC Workshop, pp. 171–179.

[10] Kakadiaris, I.A., Passalis, G., Toderici, G., Murtuza, N., Lu, Y., Karampatziakis, N., Theoharis, T., 2007. 3D face recognition in the presence of facial expressions: An annotated deformable model approach. IEEE Trans. PAMI 29 (4), 640–649.

[11] Bronstein, A.M., Bronstein, M.M., Kimmel, R., 2007. Expression-invariant representations of faces. IEEE Trans. Image Process. 16 (1), 188–197.

[12] Mpiperis, I., Malassiotis, S., Strintzis, M.G., 2007. 3-D face recognition with the geodesic polar representation. IEEE Trans. Inform. Forensics Security 2 (3), 537–547.

[13] Pears, N., Heseltine, T., 2006. Isoradius contours: New representations and techniques for 3D face registration and matching. In: Proc. 3rd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'06).

[14] Chenghua Xu, Yunhong Wang, Tieniu Tan and Long Quan, Automatic 3D Face Recognition Combining Global Geometric Features with Local Shape Variation Information, Proc. The 6th IEEE International Conference on Automatic Face and Gesture Recognition (FG), pp.308-313, 2004.

[15] Besl, P.J., Mckay, N.D., 1992. A method for registration of 3-D shapes. IEEE Trans. PAMI 14 (2), 239–256.

[16] Bock, H. H. Diday E. (Eds) : "Analysis of Symbolic Data", Springer Verlag (2000).

[17] Carlo N. Lauro and Francesco Palumbo, "Principal Component Analysis of Interval Data: a Symbolic Data Analysis Approach", Computational Statistics, Vol.15 n.1 (2000) pp.73-87.

[18] Xu, C., Wang, Y., Tan, T., Quan, L., 2006b. A robust method for detecting nose on 3D point cloud. Pattern Recognition Lett. 27 (13), 1487–1497.

[19] Xu, C., Wang, Y., Tan, T., Quan, L., 2004. 3D face recognition based on G-H shape variation. LNCS 3338. Springer, pp. 233–243.

[20] Li Li, Chenghua Xu, Wei Tang, Cheng Zhong, "3D face recognition by constructing deformation invariant image", Pattern Recognition Letters, Vol 29, pp.1596-1602 (2008).