PAPER

3D Face Recognition Using Geodesic PZM Array from a Single Model per Person

Farshid HAJATI[†], *Student Member*, Abolhasem A. RAIE^{†a)}, *Member*, and Yongsheng GAO^{††,†††}, *Nonmember*

SUMMARY For the 3D face recognition numerous methods have been proposed, but little attention has been given to the local-based representation for the texture map of the 3D models. In this paper, we propose a novel 3D face recognition approach based on locally extracted Geodesic Pseudo Zernike Moment Array (GPZMA) of the texture map when only one exemplar per person is available. In the proposed method, the function of the PZM is controlled by the geodesic deformations to tackle the problem of face recognition under the expression and pose variations. The feasibility and effectiveness investigation for the proposed method is conducted through a wide range of experiments using publicly available BU-3DFE and Bosphorus databases including samples with different expression and pose variations. The performance of the proposed method is compared with the performance of three state-of-the-art benchmark approaches. The encouraging experimental results demonstrate that the proposed method achieves much higher accuracy than the benchmarks in single-model databases.

key words: 3D face, geodesic distance, moment, PZM

1. Introduction

Face recognition is a challenging task because of the variety in expression, age, pose, illumination, and occlusion [1]. 2D face algorithms provide good performance under controlled conditions, but their performance reduces in the presence of condition variations. Recently, 3D information of the face has been used to tackle this problem. 3D face recognition approaches can be divided into two categories: range approaches and range-texture approaches. The range approaches only use the features extracted from the range (depth) maps for recognition. Nagamine et al. [2] aligned face range maps and three curves at which the face surface has an intersection with vertical and horizontal planes and a cylinder for matching. Beumier and Acheroy [3] used local curvature values along the profile curves to compare the profile range maps. Tanaka et al. [4] deals with the face recognition problem as a 3D shape recognition problem of non-rigid surfaces. Their method is based on the surface curvature information. They adopted an Extended Gaussian Image (EGI) as a mediate feature after a curvature-based segmentation. Chua et al. [5] used the point signature in recognizing

frontal faces with expression variations. They extracted the rigid parts of a face to deal with different facial expressions. Heshner et al. [6] applied PCA to range maps and estimated the probability models for the coefficients.

The range-texture approaches are based on the algorithms using both range and texture maps for recognition. Chang et al. [7] demonstrated that the rank-1 recognition rates on both texture and range maps are alike. They also showed that the combination of texture and range maps improves the recognition rate. Malassiotis and Srinivas [8] used a method for face recognition using both texture and range maps. They used the range map to compensate pose direction in the corresponding texture map. They compensated the pose by comparing the range map with a range model by using the ICP algorithm [9]. Bronstein et al. [10], [11] assumed that the facial expressions can be modeled as isometries of facial surfaces. They proposed a representation of faces called canonical formes. The canonical formes were constructed by geodesic distances and Multi-Dimensional Scaling (MDS). Mpiperis et al. [12] introduced a Geodesic Polar Representation for 3D face recognition. They used a geodesic-based isometric mapping to provide warped texture maps and apply PCA technique for classification.

Different methods have been used to address the face recognition based on 3D models, but most of them have focused on only using the range map. In cases that they used the texture map, they used it to improve the recognition rate on range map. However, the texture map plays an important role in the face recognition scenario, especially when the range maps of the faces in the gallery are similar. Moreover, all proposed methods have used the global features of the texture map which is sensitive to pose and expression changes, while local descriptors have gained much attention in face community for their robustness to face variations. Penev and Atick et al. [13] proposed a second-order statistical Local Feature Analysis (LFA) to extract the local features. Elastic Bunch Graph Matching (EBGM) [14] presented a local representation for the face using a group of Gabor coefficients. Heisele et al. [15] investigated the feasibility of the patch based face recognition and outperformed the holistic approaches. Local Binary Pattern (LBP) as a local descriptor has been used as a texture descriptor [16], facial expression analysis [17], background modeling [18], and face recognition [19]. Zhang et al. [20] proposed the Local Derivative Pattern (LDP) as a high-order local pattern descriptor for face recognition based on the local derivative

Manuscript received October 12, 2010.

Manuscript revised February 8, 2011.

[†]The authors are with the Faculty of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran.

^{††}The author is with the School of Engineering, Griffith University, QLD 4111, Australia.

^{†††}The author is with the Queensland Research Lab, National ICT Australia, QLD 4207, Australia.

a) E-mail: raie@aut.ac.ir

DOI: 10.1587/transinf.E94.D.1488

variations. Local texture representation, however, still remains an unexplored area in 3D face recognition field.

We present a novel 3D face recognition approach using Geodesic Pseudo Zernike Moment Array (GPZMA). Geodesic distance is a 3D feature descriptor that can analyze the deformations on face surface caused by expression and rotation variations [10], [11]. The Pseudo Zernike Moment (PZM) is a powerful texture descriptor in 2D images [26]. In this paper, we propose a novel 3D face recognition by combining the geodesic distance and Pseudo Zernike Moments in which the function of the Zernike Moments is controlled by the computed face geodesic distance making PZMs more robust to face variations than the traditional methods that use the Pseudo Zernike polynomials and geodesic distance, separately. Unlike the other methods in this field, the proposed method is applied on the texture information to extract local feature vectors with deformations controlled by the range information. The proposed method is evaluated by the publicly available BU-3DFE face database [21] and Bosphorus face database [22] both of which contain faces under various situations including different expression and pose changes. The effectiveness and performance of the proposed method is compared with the effectiveness and performance of the state-of-the-art approaches. Encouraging experimental results indicate that the proposed method provides a novel solution for 3D face recognition.

The organization of the paper is as follows: Section 2 presents the proposed Geodesic Pseudo Zernike Moment Array. In Sect. 3, the proposed method is evaluated and compared with benchmark approaches. Finally, Sect. 4 presents the conclusion and the future work.

2. Proposed Geodesic Pseudo Zernike Moment Array

The overall framework of the proposed method is illustrated in Fig. 1. In this method, we create the corresponding range and texture maps from an input 3D model. Using the range map, we compute the geodesic distance from a reference point for all face points. Based on the computed geodesic distances, we transform the texture map to a Transformed Texture Map. In the next step, the transformed texture map is partitioned into equal-sized square patches and a PZM descriptor is applied to extract a feature vector for each patch. By concatenating feature vectors, we build a feature array for the subject. Finally, the input subject is identified by measuring the dissimilarity between the query and all gallery models. The following subsections will describe the algorithm in detail.

2.1 Geodesic Distance Computation

Let S be a surface, P and Q two points on S as shown in Fig. 2. There are an infinite number of curves that belong to S and connect P with Q . The curve on the surface S with the minimum length is called the Geodesic Path between two points and its length is called the Geodesic Distance. The numerical computation of the geodesic distance involves the solving of Eikonal equation [23]

$$|\nabla T(x, y, z)| = 1 \quad (1)$$

on the surface, where $T(x, y, z)$ is the geodesic distance of the surface point $Q(x, y, z)$ from a reference point. The main approach to solve the Eikonal equation in a 3D space is the Fast Marching Method on Triangulated Domains

Fig. 1 Framework of the proposed method.

Fig. 2 Geodesic path between P and Q.

Fig. 3 An arbitrary face 3D model and its range and texture maps. (a) 3D model, (b) range map, (c) texture map.

(FMTD) [24].

In this paper, we compute the geodesic distances on the face surface using the range map extracted from the face 3D models. If we represent the range map as $z = f^r(x, y)$, in which the superscript r stands for range, the Eq. (1) can be solved on the range map as

$$|\nabla T(x, y, f^r(x, y))| = 1 \quad (2)$$

where $T(x, y, f^r(x, y))$ is the geodesic distance of the point $Q(x, y, f^r(x, y))$ from a reference point. Figures 3 (a) and 3 (b) illustrate an arbitrary face 3D model and its range map, respectively.

2.2 Geodesic Texture Transform

It has been proved that the geodesic distance can be used to represent the texture and the shape in the face recognition algorithms [10]–[12], [25]. In this paper, we use the notion of face isometric assumption to find a novel texture transform which is robust to expression and pose variations. Our transform is based on the assumption that the face surface is a 2D manifold embedded in a 3D Euclidean space and we can characterize the face surface by a Riemannian metric and describe it by geodesic properties.

We represent the texture map of a 3D image as $I = f^t(x, y)$, in which the superscript t stands for texture. Figures 3 (a) and 3 (c) illustrate an arbitrary face 3D model and its texture map, respectively. For each pixel in the texture map we have a geodesic distance computed from the range map. We use the geodesic distance of the face points to establish a transform function for the texture map. The transform of the point (x, y) in the texture map is defined as

$$F^t(x_{trans}, y_{trans}) = f^t(x, y) \quad (3)$$

where

$$x_{trans} = r(x, y) \cos(\phi(x, y)) \quad (4)$$

$$y_{trans} = r(x, y) \sin(\phi(x, y)) \quad (5)$$

where $F^t(x_{trans}, y_{trans})$ denotes the transformed texture map

in which the subscript *trans* stands for transformed. x_{trans} and y_{trans} are the coordinates of the transformed texture map. $r(x, y)$ is the geodesic distance of the point $Q(x, y, f^r(x, y))$ from the reference point as

$$r(x, y) = T(x, y, f^r(x, y)) \quad (6)$$

where $T(x, y, f^r(x, y))$ is the geodesic distance computed by (2).

In (4) and (5), $\phi(x, y)$ is the angle made by the connecting line of $Q(x, y, f^r(x, y))$ and the reference point as well as the horizon line in the range map defined as

$$\phi(x, y) = \begin{cases} 0, & x = x_{RP}, y = y_{RP} \\ \tan^{-1}\left(\frac{y - y_{RP}}{x - x_{RP}}\right), & x \geq x_{RP} \\ \tan^{-1}\left(\frac{y - y_{RP}}{x - x_{RP}}\right) + \pi & x < x_{RP} \end{cases} \quad (7)$$

where x_{RP} and y_{RP} are the coordinates of the Reference Point (RP) in the range map. In this paper, we select the nose tip as the reference point and crop all transformed texture maps to the size of 160×160 in the way that the distance of the nose tip is 80 pixels from the sides, 90 and 70 pixels from the top and the bottom, respectively.

2.3 PZM Descriptor

Moment invariants have been proven to be successful in pattern recognition applications due to their power in discriminating between pattern features [26], [27]. In this paper, we use Pseudo Zernike Moments (PZM) as descriptor. Pseudo Zernike Moment (PZM) is a modification of Zernike Moments (ZM) [28] based on a different set of orthogonal polynomials that have properties analogues to Zernike polynomials [29]. The advantages of the Pseudo Zernike Moments for face recognition can be described in three aspects: 1) sensitivity to image noise, 2) information redundancy, and 3) capability for image representation. A large investigation has been done in [26] to compare Pseudo Zernike Moments in these aspects. The experiments have showed that Pseudo Zernike Moments are more robust to noise than other common moments such as Geometric Moments, Legendre Moments, Zernike Moments, Rotational Moments, and Complex Moments [26]. From the second aspect view, Pseudo Zernike Moments are orthogonal and uncorrelated moment sets. From the point of view of the third aspect, they have considered how well an image can be characterized by a small finite set of moments by using the mean-square error between an image and its reconstructed version. Experiments have showed that the Pseudo Zernike Moments have the best reconstruction performance compared to other moments [26].

The two-dimensional complex PZM of order n with repetition m of a continuous image intensity function $f(x, y)$ can be defined as

$$PZM_{n,m}(f(x, y)) = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} V_{n,m}^*(x, y) f(x, y) dx dy \quad (8)$$

where $n = 0, 1, 2, \dots$ and $|m| \leq n$ and the symbol $*$ denotes the complex conjugate.

The Pseudo Zernike polynomials $V_{n,m}(x, y)$ are defined as [26]:

$$V_{n,m}(x, y) = R_{n,m}(l) e^{jm\theta} \quad (9)$$

where $l = \sqrt{x^2 + y^2}$ is the length of the vector $\mathbf{l}$ from the origin to the pixel (x, y) and $\theta = \tan^{-1}(y/x)$ is the angle between vector $\mathbf{l}$ and x -axis.

In Pseudo Zernike polynomials, the real-valued radial polynomials $R_{n,m}(l)$ are defined as [26]:

$$R_{n,m}(l) = \sum_{s=0}^{n-|m|} (-1)^s \frac{(2n+1-s)!}{s!(n-|m|-s)!(n+|m|+1-s)!} l^{(n-s)} \quad (10)$$

where $n = 0, 1, 2, \dots$ and m takes positive and negative integer values subject to $|m| \leq n$.

In this paper, we use PZM as a local descriptor in the proposed face recognition system. This descriptor is applied on the transformed texture map to extract feature vectors. For this purpose, we first partition the transformed texture map to equal-sized square patches in a non-overlapping way. Assume that the size of the transformed texture map is $N \times N$ and the size of each patch is $W \times W$. In this case, the number of patches for the transformed texture map is $(N/W)^2$. We indicate all points in a patch with u and v indexes which are integers ranging from 1 to N/W and mathematically can be found for each point as

$$u = \left\lfloor \frac{x_{trans}}{W} \right\rfloor + 1 \quad 0 \leq x_{trans} < N \quad (11)$$

$$v = \left\lfloor \frac{y_{trans}}{W} \right\rfloor + 1 \quad 0 \leq y_{trans} < N \quad (12)$$

which x_{trans} and y_{trans} are coordinates of the transformed texture map. The symbol $\lfloor \cdot \rfloor$ denotes the floor function. Figure 4 illustrates an example of the partitioning a transformed texture map into four patches.

Given a transformed texture map, $F^t(x_{trans}, y_{trans})$, the (u, v) th patch in the transformed texture map $F^t_{u,v}(x_{trans}^{uv}, y_{trans}^{uv})$ is defined as

$$F^t(x_{trans}, y_{trans}) = F^t_{u,v}(W(u-1) + x_{trans}^{uv}, W(v-1) + y_{trans}^{uv}) \quad (13)$$

where x_{trans}^{uv} and y_{trans}^{uv} are coordinates of the (u, v) th patch.

Fig. 4 An example of the partitioning a transformed texture map into four patches.

The origin of the patch coordinates system is located at the bottom-left corner of each patch.

By using (13) and (8), the PZM equation for the (u, v) th patch in the transformed texture map is derived as

$$PZM_{n,m}^{uv}(F^t(x_{trans}, y_{trans})) = \frac{n+1}{\pi} \iint_{x_{trans}^{uv^2} + y_{trans}^{uv^2} \leq 1} V_{n,m}^*(x_{trans}^{uv}, y_{trans}^{uv}) F^t_{u,v}(x_{trans}^{uv}, y_{trans}^{uv}) dx_{trans}^{uv} dy_{trans}^{uv} \quad (14)$$

where u and v are integers ranging from 1 to N/W .

To compute the PZM for each patch, the texture of the patch is normalized into a unit circle. There are two possible approaches for this normalization. One is the normalization of the patch in the way that the unit circle is inside the patch. In this way, the information of the pixels located outside the unit circle will be lost. The second approach, we use in this paper, is the normalization of the patch in the way that the entire patch is bounded inside the unit circle. This approach ensures that there is no pixel loss during the PZM computation. Figure 5 illustrates an example of the normalization of an arbitrary patch into the unit circle.

Let assume the size of a patch is $W \times W$. The (u, v) th patch of the transformed texture map is mapped into the unit circle as

$$x_{trans,i}^{uv} = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{W-1}i \quad i = 0, 1, \dots, W-1 \quad (15)$$

$$y_{trans,j}^{uv} = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{W-1}j \quad j = 0, 1, \dots, W-1 \quad (16)$$

where $x_{trans,i}^{uv}$ and $y_{trans,j}^{uv}$ are coordinates in the normalized (u, v) th patch.

The Eq. (14) is the continuous form of the PZM. By using (15) and (16), we can derive a discrete form of PZM for a digital transformed texture map. Given the (u, v) th patch in the digital transformed texture map $F^t(x_{trans}, y_{trans})$, the discrete PZM with order n and repetition m is derived as

Fig. 5 Normalization of an arbitrary patch into the unit circle.

$$\begin{aligned}
& PZM_{n,m}^{uv} \left(F^t(x_{trans}, y_{trans}) \right) \\
&= \frac{n+1}{\pi \lambda(W)} \sum_{i=0}^{W-1} \sum_{j=0}^{W-1} V_{n,m}^* \left(x_{trans,i}^{uv}, y_{trans,j}^{uv} \right) \\
& \quad F_{u,v}^t \left(x_{trans,i}^{uv}, y_{trans,j}^{uv} \right)
\end{aligned} \quad (17)$$

where the normalization factor $\lambda(W)$ is the ratio between the number of pixels in the patch before and after mapping into the unit circle. This normalization factor for our mapping is $W^2/2$.

In this paper, the magnitudes of PZMs of order $n_{\min}$ up to $n_{\max}$ with repetition $m \geq 0$ will be considered as moment features. $n_{\min}$ and $n_{\max}$ will be defined in Sect. 3.1 experimentally. The feature vector for the (u, v) th patch is defined as

$$\begin{aligned}
PZM_{n_{\min}, n_{\max}}^{uv} &= \left\{ \left| PZM_{a,b}^{uv} \left(F^t(x_{trans}, y_{trans}) \right) \right| \right\} \\
& \quad a = n_{\min}, \dots, n_{\max} \& b = 0, \dots, a
\end{aligned} \quad (18)$$

where the symbol $|\cdot|$ denotes the absolute value operator. The output of the PZM descriptor is a feature vector for each patch in the transformed texture map. In the next section, we use these feature vectors to calculate the dissimilarity between two subjects.

2.4 Dissimilarity Measurement

In this section, we calculate the distance of a given query transformed texture map and a model transformed texture map in the gallery using extracted PZMs. A Patch Distance Vector (PDV) is the distance vector between the (u, v) th patch of the transformed texture maps in the query and the model as

$$PDV_{u,v} = \left\{ \left[PZM_{n_{\min}, n_{\max}}^{uv} \right]_f - \left[PZM_{n_{\min}, n_{\max}}^{uv} \right]_g \right\} \quad (19)$$

where f and g are the transformed texture maps of the query and the model, respectively.

In a boundary patch, the patch located on the face border, some pixels of the patch are set to zero based on the shape of the texture (see Fig. 6). This zero valued pixels will affect the computed Zernike Moments. Therefore, the computed distance between two patches in the query and model transformed texture map will be affected by the shape of the texture in the boundary patches. To decrease this effect on the extracted moments, an adaptively weighing technique based on the textural area of each patch is used to weigh the moment features of each patch. For this purpose, we define the weight of each patch based on the patch textural area in both the query transformed texture map and the corresponding patch textural area in the model transformed texture map. To decrease the effect of the boundary patches in the final distance function, we map both patches into a unit circle and compute the proportion of the overlapping area of two patches to the whole area as a weight $0 \leq \eta_{u,v} \leq 1$:

$$\eta_{u,v} = \frac{S_{u,v}^Q \cap S_{u,v}^M}{2} \quad (20)$$

Fig. 6 Computation of $S_{u,v}^Q \cap S_{u,v}^M$ for two arbitrary patches in the query and the model transformed texture maps.

where $\eta_{u,v}$ is the weight for the (u, v) th patch, $S_{u,v}^Q$ and $S_{u,v}^M$ are the textural areas of the (u, v) th patch of the transformed texture maps of the query and model in the unit circle, respectively. The superscripts Q and M stand for query and model, respectively. The symbol $\cap$ denotes the overlapping operator. Figure 6 illustrates the computation of $S_{u,v}^Q \cap S_{u,v}^M$ for two arbitrary patches in the query and model transformed texture maps. In this weighing technique, a smaller weight ($\eta_{u,v} < 1$) is dedicated for the boundary patches, while the weight is $\eta_{u,v} = 1$ for non-boundary patches.

By using $PDV_{u,v}$ and $\eta_{u,v}$, we compute the distance between the query transformed texture map (f) and the model transformed texture map (g) as

$$D(f, g) = \left\| \left\{ \eta_{u,v} \cdot PDV_{u,v} \mid u, v = 1, \dots, \frac{N}{W} \right\} \right\| \quad (21)$$

where the symbol $\|\cdot\|$ denotes the mathematical multiplication operator.

For a given query, we compute the distance to all models in the gallery using (21) and consider the model with the minimum distance D as the correct return.

3. Experimental Results

In order to evaluate the proposed approach, an extensive experimental investigation is conducted, covering face recognition under different pose and expression changes. For the experiments, we use the BU-3DFE database [21] which contains 3D face models under all different expression

changes and Bosphorus database [22] which contains 2.5D face scans from different pose variations. The BU-3DFE database contains 100 subjects (56 females and 44 males) and each subject has seven expressions. With the exception of the neutral expression, each of the six expressions (angriness, disgust, fear, happiness, sadness, and surprise) includes four levels of intensity. The Bosphorus database is tested in this paper to measure the performance of the proposed method under the pose changes. It contains over 5200 2.5D scans from 105 subjects (61 females and 44 males) in different poses. The performance of the proposed method is compared with that of three benchmark approaches: 1) the method proposed by Bronstein for expression-invariant face recognition (Canonical Image Representation) [10], 2) the method proposed by Mpiperis for expression-invariant face recognition (Geodesic Polar Representation) [12], and 3) the method proposed by Malassiotis and Strintzis [8] for pose-invariant face recognition. In all experiments, a preprocessing of face localization was applied.

3.1 Investigation and Determination of Parameters

As described in Sect. 2.3, the Pseudo Zernike polynomials have two parameters m and n involved in the proposed method. $n \geq 0$ is the order and $|m| \leq n$ is the repetition of the PZMs. In this section, we investigate and determine these two parameters accompanying with the size of patches (W).

In the proposed method, because of $R_{n,-m}(r) = R_{n,m}(r)$ we use all $m = 0, 1, \dots, n$ as in (18). Therefore, we have all repetitions in the result feature vector. However, we use Pseudo Zernike polynomials of order $n_{\min}$ up to $n_{\max}$. In theory, the PZM with order zero represents the DC component of the image and does not have personal identity information. By setting $n = 0$ in (17), we have

$$\begin{aligned} PZM_{0,0}^{uv}(F(x_{trans}, y_{trans})) \\ &= \frac{2}{\pi W^2} \sum_{i=0}^{W-1} \sum_{j=0}^{W-1} V_{0,0}^* (x_{trans,i}^{uv}, y_{trans,j}^{uv}) \\ &\quad F_{u,v}^t (x_{trans,i}^{uv}, y_{trans,j}^{uv}) \\ &= \frac{2}{\pi} F_{u,v}^t (x_{trans,i}^{uv}, y_{trans,j}^{uv}) \end{aligned} \quad (22)$$

where $F_{u,v}^t (x_{trans,i}^{uv}, y_{trans,j}^{uv})$ is the mean value of pixels in the (u, v) th patch of the normalized transformed texture map. Hence, we have set $n_{\min} = 1$.

To determine $n_{\max}$ and W , an experimental investigation on recognition rate is conducted on a training dataset. For this purpose, we created a training dataset from the BU-3DFE database [21]. For each subject, we selected two faces: one is the neutral face and another is one of the six expressions with one of four expression levels selected randomly. In the training dataset, all neutral faces are selected as the gallery and others are used as the probe. Using the training dataset, an experimental investigation on recognition accuracy is conducted with different values of W (10,

Fig. 7 The effect of W and $n_{\max}$ on the recognition rate.

16, 20, 32, 40, 80, and 160) and $n_{\max}$ (from 1 up to 10). The result is displayed in Fig. 7.

It can be observed that through reducing the patch size, system accuracy continuously increases and reaches the highest rate when $W=16$, then decreases when $W=10$. It is encouraging to see that the required $n_{\max}$ (*i.e.* the value of $n_{\max}$ when the curve starts to remain flat) continuously decreases when W decreases. In the rest of the experiments we choose $W=16$ as the patch size, and $n_{\max}=3$ as the maximum order of PZM.

3.2 Face Recognition under Facial Expression Changes

Experiments were conducted on BU-3DFE database [21] to evaluate the effects of different facial expressions (angriness, disgust, fear, happiness, sadness, and surprise) on the proposed approach. Figure 8 illustrates an example of the facial expressions in the BU-3DFE database. To have more reliable comparison with the results of benchmarks reported in [12], the same experimental strategy used in [12] is used in our study. The performance is measured in term of the rank-1 recognition rate and equal error rate (ERR) [30]. For each person, a single-neutral model is used as representative in the gallery, while the rest images depicting expressions are used as probe. In total we have six probe sets for the recognition rate under six different face expression changes.

The rank-1 recognition rates and equal error rates (EERs) of the proposed algorithm for each type of expression changes and for the whole database are tabulated in Table 1. For comparison, the rank-1 recognition rates, EERs, CMC curves, and ROC curves for benchmarks are reported from [12]. As can be seen, the angry expression has the best accuracy (90%), while the disgust expression has the worst accuracy (78%). The overall accuracy for the proposed method (82.2%) is higher than the benchmark approaches (80.3% and 77.2% for Geodesic Polar Representation and Canonical Image Representation, respectively). Moreover, the equal error rate (EER) of the proposed method is 6.7% compared to 9.8% and 15.4% for Geodesic Polar Representation and Canonical Image Representation, respectively. Two different recognition rates on two different

Fig. 8 Different types and levels of expression changes in the BU-3DFE database.

Table 1 Performance comparison under facial expression changes. The results marked by * are from [12].

Expression Considered	Proposed Method	*Geodesic Polar Representation	*Canonical Image Representation
Angriiness	90	N/A	N/A
Disgust	78	N/A	N/A
Fear	80	N/A	N/A
Happiness	82	N/A	N/A
Sadness	82	N/A	N/A
Surprise	81	N/A	N/A
Overall	82.2	80.3	77.2
EER	6.7	9.8	15.4

databases have been reported in [12]. The first database is the BU-3DFE database which is prepared by State University of New York at Binghamton. This database is publicly available. The second database is a self-prepared 3D face database which is not publicly available for other researchers. The recognition rate of the Geodesic Polar Representation method is 80.3% and 90.4% on two databases, respectively. While acknowledging the finding of the second database, the results of our proposed algorithm i.e. 82.2%, due to the accessibility of the first one, is just compared with the former which is 80.3%. The CMC and ROC curves of the proposed method together with those of the two benchmark methods are given in Fig. 9 and Fig. 10. Note that the

Fig. 9 Recognition rates obtained under facial expression changes.

Fig. 10 Receiver operating characteristics for the proposed method, geodesic polar representation, and canonical image representation.

recognition rate of the proposed method is always higher than the benchmark methods.

3.3 Face Recognition under Pose Rotations

In this section, we evaluate the robustness of the proposed method under different face rotations. In this experiment, we use the Bosphorus database [22] which contains 2.5D face scans in different rotations. Figure 11 illustrates an arbitrary subject from the database in different pose rotations. The parameters of the algorithm (*i.e.* the minimum of the PZMs, the maximum of the PZMs, and the size of the patch) are those selected in Sect. 3.1 using the training data set.

The rank-1 recognition rates of the proposed algorithm for all rotation angles are tabulated in Table 2. For comparison, the rank-1 recognition rates of the method proposed by Masassiotis and Srinivas [8] are used as the benchmark. As can be seen, the proposed method has much higher accuracy compared to the benchmark. The proposed method has 87.6% recognition rate in $+10^\circ$ right rotation, while the benchmark has the recognition rate of 65.7%. In $+20^\circ$ right rotation, the recognition rate is 84.7% and 52.4% for the proposed method and the benchmark, respectively. In $+30^\circ$ right rotation, the recognition rate is 74.3% and 42.9% for

Fig. 11 Different types of rotations in the Bosphorus face database.

Table 2 Performance comparison under pose rotations.

Rotation Considered	Proposed Method	Malassiotis and Strintzis Method [8]
+10° Right	87.6	65.7
+20° Right	84.7	52.4
+30° Right	74.3	42.9
+45° Right	34.3	9.5
-45° Left	34.3	21
Slight Upwards	83.8	81.9
Upwards	74.3	69.5
Slight Downwards	85.7	84.8
Downwards	69.5	68.6
Right-Upwards	59	31.4
Right-Downwards	29.5	21.9
Overall	65.2	50

the proposed method and the benchmark, respectively. The overall recognition rate in the proposed method is much higher than the recognition rate in the benchmark (65.2% compared to 50%). It is obvious that if we set the parameters (*i.e.* the maximum of the PZMs and the size of the patch) for the Bosphorus database, the results in this section will be higher than those reported here.

4. Conclusion and Further Research

In this paper, a novel 3D face recognition method has been proposed called Geodesic Pseudo Zernike Moment Array (GPZMA). GPZMA is particularly designed to handle the variations in face surface due to the expression and pose in single-model databases. GPZMA uses the geodesic distance between face surface points to transform the texture map of the 3D model and to extract moments. The Pseudo Zernike Moments controlled by geodesic distance are adapted to be used in the proposed face recognition approach as local descriptor. The idea behind the work is that better results can be obtained if local texture information of 3D models are used instead of global information. The algorithm has been evaluated and compared with three state-of-the-

art benchmarks. It is very encouraging to find that the proposed method performs consistently superior to the benchmarks under the expression and poses variations. The presented experimental results indicate an overall improvement of 1.9% and 5% for expression variations and 15.2% for pose variations compared to the benchmarks.

Because the main novelty of this paper is in the recognition phase of a 3D face recognition system, the positioning of the reference point (nose tip) has been done manually in the same strategy as in [10], [12]. Changing the reference point may change the geodesic distance of the face surface points. The effect of automatically positioning accuracy of the reference point on the overall performance of the complete system might be investigated as a future research project.

References

- [1] A.F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognit. Lett.*, vol.28, no.14, pp.1885–1906, 2007.
- [2] T. Nagamine, T. Uemura, and I. Masuda, "3D facial image analysis for human identification," *Proc. International Conference on Pattern Recognition*, pp.324–327, Hague, The Netherlands, Aug. 1992.
- [3] C. Beumier and M. Achery, "Automatic 3D face authentication," *Image Vis. Comput.*, vol.18, no.4, pp.315–321, 2000.
- [4] H.T. Tanaka, M. Ikeda, and H. Chiaki, "Curvature-based face surface recognition using spherical correlation-principal directions for curved object recognition," *Proc. 3rd International Conference on Face and Gesture Recognition*, pp.372–377, Nara, Japan, April 1998.
- [5] C. Chua, F. Han, and Y. Ho, "3D human face recognition using point signature," *Proc. 4th International Conference on Automatic Face and Gesture Recognition*, pp.233–238, Grenoble, France, March 2000.
- [6] C. Heshner, A. Srivastava, and G. Erlebacher, "Principal component analysis of range images for facial recognition," *Proc. International Conference on Imaging Science, Systems, and Technology*, pp.62–68, Las Vegas, USA, 2002.
- [7] K.I. Chang, K.W. Bowyer, and P.J. Flynn, "Multi-modal 2D and 3D biometrics for face recognition," *Proc. IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, pp.187–195, Nice, France, Oct. 2003.
- [8] S. Malassiotis and M.G. Strintzis, "Robust face recognition using 2D and 3D data: Pose and illumination compensation," *Pattern Recognit.*, vol.38, no.12, pp.2537–2548, 2005.
- [9] P.J. Besl and N.D. McKay, "A method for registration of 3D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.14, no.2, pp.239–256, 1992.
- [10] A.M. Bronstein, M.M. Bronstein, and R. Kimmel, "Expression-invariant representations of faces," *IEEE Trans. Image Process.*, vol.16, no.1, pp.188–197, 2007.
- [11] A.M. Bronstein, M.M. Bronstein, R. Kimmel, and A. Spira, "Face recognition from facial surface metric," *Proc. 8th European Conference on Computer Vision*, Prague, pp.225–237, Czech Republic, May 2004.
- [12] I. Mpiperis, S. Malassiotis, and M.G. Strintzis, "3-D face recognition with the geodesic polar representation," *IEEE Trans. Information Forensics and Security*, vol.2, no.3, pp.537–547, 2007.
- [13] P.S. Penev and J.J. Atick, "Local feature analysis: A general statistical theory for object representation," *Netw., Comput. in Neural Syst.*, vol.7, no.3, pp.447–500, 1996.
- [14] L. Wiskott, J. Fellous, N. Kruger, and C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal.*

- Mach. Intell., vol.19, no.7, pp.775–779, 1997.
- [15] B. Heisele, T. Serre, and T. Poggio, “A component-based framework for face detection and identification,” *Int. J. Comput. Vis.*, vol.74, no.2, pp.167–181, 2007.
 - [16] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.7, pp.971–987, 2002.
 - [17] G. Zhao and M. Pietikainen, “Dynamic texture recognition using local binary patterns with an application to facial expressions,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.29, no.6, pp.915–928, 2007.
 - [18] M. Heikkila and M. Pietikainen, “A texture-based method for modeling the background and detecting moving objects,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.28, no.4, pp.657–662, 2006.
 - [19] T. Ahonen, A. Hadid, and M. Pietikainen, “Face description with local binary patterns: Application to face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.28, no.12, pp.2037–2041, 2006.
 - [20] B. Zhang, Y. Gao, S. Zhao, and J. Liu, “Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor,” *IEEE Trans. Image Process.*, vol.19, no.2, pp.533–544, 2010.
 - [21] L. Yin, X. Wei, Y. Sun, J. Wang, and M.J. Rosato, “A 3D facial expression database for facial behavior research,” *Proc. 7th International Conference on Automatic Face and Gesture Recognition (FGR)*, pp.211–216, Southampton, UK, April 2006.
 - [22] A. Savran, N. Alyuz, H. Dibeklioglu, O. Celiktutan, B. Gokberk, B. Sankur, and L. Akarun, “Bosphorus database for 3d face analysis,” *Proc. First COST 2101 Workshop on Biometrics and Identity Management (BIOID)*, pp.47–56, Denmark, May 2008.
 - [23] Y.R. Tsai, L. Cheng, S. Osher, and H. Zhao, “Fast sweeping algorithms for a class of hamilton-jacobi equations,” *SIAM J. Numer. Anal.*, vol.41, no.2, pp.673–694, 2003.
 - [24] R. Kimmel and J.A. Sethian, “Computing geodesic paths on manifolds,” *Proc. National Academy Sciences*, vol.95, no.15, pp.8431–8435, 1998.
 - [25] A.M. Bronstein, M.M. Bronstein, and R. Kimmel, “Three-dimensional face recognition,” *Int. J. Comput. Vis.*, vol.64, no.1, pp.5–30, 2005.
 - [26] C.H. Teh and R.T. Chin, “On image analysis by the methods of moments,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.10, no.4, pp.496–513, 1988.
 - [27] A. Nabatchian, E. Abdel-Raheem, and M. Ahmadi, “Human face recognition using different moment invariants: A comparative study,” *Proc. Congress on Image and Signal Processing*, pp.661–666, Sayna, China, May 2008.
 - [28] M.R. Teague, “Image analysis via the general theory of moments,” *J. Optical Society of America*, vol.70, no.8, pp.920–930, 1980.
 - [29] F. Zernike, “Beugungstheorie des schneidenverfahrens und seiner verbesserten form, der phasenkontrastmethode,” *Physica*, vol.1, no.1, pp.689–704, 1934.
 - [30] S.A. Rizvi, P.J. Phillips, and H. Moon, “The FERET verification testing protocol for face recognition algorithms,” *Proc. 3rd International Conference on Face and Gesture Recognition*, pp.48–53, Nara, Japan, April 1998.

tion, computer vision, neural networks, and face detection and recognition.

Farshid Hajati received the B.Sc. degree in electronic engineering from Khaje Nasir University of Technology, Iran in 2003 and M.Sc. degree in electronic engineering from Amirkabir University of Technology, Iran in 2006. From 2006, he has been a Ph.D. student in electronic engineering at Amirkabir University of Technology. From 2009 to 2010 he was a visiting scholar with the School of Engineering, Griffith University, Australia. His research interests include digital image processing, pattern recognition,

Abolghasem A. Raie received the B.Sc. degree in electronic engineering from Sharif University of Technology, Iran in 1973 and M.Sc. and Ph.D. degrees in electronic engineering from University of Minnesota, USA in 1979 and 1982, respectively. From 1983, he has been a member of the Faculty of Electrical Engineering at AmirKabir University of Technology in Iran. His research interests are algorithm design, performance analysis, machine vision, sensor fusion, and mobile robots navigation.

nition, biometrics, biosecurity, image retrieval, computer vision, and pattern recognition.

Yongsheng Gao received the B.Sc. and M.Sc. degrees in electronic engineering from Zhejiang University, China, in 1985 and 1988, respectively, and the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore. Currently, he is an Associate Professor with the Griffith School of Engineering, Griffith University, Australia. He is also with National ICT Australia, Queensland Research Laboratory, leading the Biosecurity group. His research interests include face recognition,