J. Cent. South Univ. (2015) 22: 1817−1825
DOI: 10.1007/s11771-015-2700-x

Springer

Face recognition using SIFT features under 3D meshes

ZHANG Cheng(张诚)[1], GU Yu-zhang(谷宇章)[1], HU Ke-li(胡珂立)[2], WANG Ying-guan(王营冠)[1]

1. Key Laboratory of Wireless Sensor Network & Communication,
Shanghai Institute of Microsystem and Information Technology,
Chinese Academy of Sciences, Shanghai 201800, China;
2. Department of Computer Science and Engineering, Shaoxing University, Shaoxing 312000, China

Abstract: Expression, occlusion, and pose variations are three main challenges for 3D face recognition. A novel method is presented to address 3D face recognition using scale-invariant feature transform (SIFT) features on 3D meshes. After preprocessing, shape index extrema on the 3D facial surface are selected as keypoints in the difference scale space and the unstable keypoints are removed after two screening steps. Then, a local coordinate system for each keypoint is established by principal component analysis (PCA). Next, two local geometric features are extracted around each keypoint through the local coordinate system. Additionally, the features are augmented by the symmetrization according to the approximate left-right symmetry in human face. The proposed method is evaluated on the Bosphorus, BU-3DFE, and Gavab databases, respectively. Good results are achieved on these three datasets. As a result, the proposed method proves robust to facial expression variations, partial external occlusions and large pose changes.

Key words: 3D face recognition; scale-invariant feature transform (SIFT); expression; occlusion; large pose changes; 3D meshes

1 Introduction

Biometrics authentication is used as a form of identification and access control in computer science. Human face has unique advantages over other biometrics for applications in scenarios in which acquiring fingerprint and iris data is unrealistic or undesirable due to problems of social acceptance [1], since it is natural, nonintrusive and contactless [2]. Face recognition has received a rapid development over the past decades based on 2D images. But one of obvious drawbacks of 2D face recognition is that it is sensitive to illumination variation [3−4]. With the development of 3D information acquisition technology, it is likely and convenient to acquire the 3D data of an object. The 3D data contain shape and geometry information, so face recognition with 3D scans can be resistant to illumination changes. Hence, 3D face recognition has been getting more and more attention by researchers in recent years.

However, 3D face recognition also has some challenges. First, expressions, such as happiness, surprise and anger, cause non-rigid deformations of facial muscles. So, facial expressions increase the intra-class difference of the same object and affect the performance of face recognition. Second, since external partial occlusions obscure partial facial surface, it introduces noise and misses some useful information which increases the difficulty for recognition. Third, large posture changes cause self-occlusion and lead to the incompleteness of facial scans. Hence, 3D face recognition is still posture dependent. Therefore, for the purpose of identification in real case scenarios, a good 3D face recognition system should be required to tackle with these problems.

While some 3D face recognition methods take advantage of 2D face images in combination with 3D data, the focus of this work is on using only the 3D data for recognition. In this work, it is principally concerned with extracting robust and repeatable features with the facial expressions, pose variations and partial occlusions. After basic preprocessing, e.g. spike removal and hole filling, a difference scale space is constructed to search salient points in the facial scans. Keypoints on the facial scans are detected as shape index extrema in this scale space. Then, two procedures are applied to eliminating the volatile keypoints. Next, a local coordinate system is established by principal component analysis (PCA) method for each keypoint. One distance feature and one angle feature are extracted around each keypoint based

Foundation item: Project(XDA06020300) supported by the "Strategic Priority Research Program" of the Chinese Academy of Sciences; Project(12511501700) supported by the Research on the Key Technology of Internet of Things for Urban Community Safety Based on Video Sensor networks
Received date: 2014−04−04; **Accepted date:** 2014−10−30
Corresponding author: ZHANG Cheng, PhD; Tel: +86−21−69075501; E-mail: mjcheng@mail.sim.ac.cn

on the local coordinate system. Since the local coordinate system is independent of spatial position of facial scans, these two features are invariant to changes of spatial location of facial scans. Moreover, to reduce the problem of loss of information caused by large pose variations and external occlusions, the features increase by symmetrization on the basis of the left-right symmetry in human face.

2 Related work

3D face recognition techniques can be classified into two categories according to the matching strategies, namely, spatial direct matching and feature-based matching techniques. Moreover, the feature-based matching techniques can be divided into the local and holistic subtypes further.

The spatial direct matching techniques make point clouds as input and can match two point clouds without feature extraction explicitly. Iterative closest point (ICP) and Hausdorff distance algorithm are the two mainstream spatial direct matching techniques. ICP was first proposed by BESL and MCKAY [5], which was used for alignment between 3D data. Several works have explored ICP or its modified versions on 3D facial surface to match faces [6−11]. Hausdorff distance can be used to determine the degree of resemblance between two objects that are superimposed on one another [12]. ACHERMANN and BUNKE [13] firstly measured the similarity between 3D facial scans with Hausdorff distance. LEE and SHIM [14] proposed a person verification approach based on depth-weighted Hausdorff distance by using the surface curvatures of the human face. RUSS et al [15] presented a Hausdorff distance matching algorithm based on the 2D range image.

In the field of 3D face recognition, the local features contain local descriptor, curve features, curvature and some other local geometric and statistic features. Spin image [16−18] and point signature [19−22] are two main local descriptors used by researchers. Curve features extract discriminative surface curves for facial representation. Three types of curves on the facial surface, such as vertical, horizontal, and circular curves, were used as the distinguishing surface features in Ref. [23]. BEUMIER and ACHEROY [24] utilized the central profile with maximal protrusion and two parallel profiles for face recognition. SAMIR et al [25] used unions of the level curves of the depth function to describe 3D facial surfaces. Curvature features include Gaussian curvature, mean curvature, shape index and so on. They have been employed for 3D face recognition in some works [26−28].

The holistic feature-based matching techniques extract the global features of facial scans and can mainly be grouped into three classes. First, the 3D face is represented as 2D depth image and then 2D face recognition algorithms are utilized to recognize faces. Second, the 3D face is mapped as extended Gaussian image (EGI) [29] and then EGI is used for matching faces. The third type of method transforms 3D face into another 3D form which is applied for face recognition. BRONSTEIN et al [30] constructed expression-invariant representations of faces based on the bending-invariant canonical forms approach.

3 Proposed approach

SIFT is used to extract and describe local distinctive invariant features in images, which was first proposed in 1999 [31] and improved in 2004 [32] by LOWE. It has a wide range of application including object recognition [32], robot navigation and tracking [33], image stitching [34], and video tracking [35]. Due to the good performance of SIFT in 2D computer vision, SIFT algorithm has been extended to higher dimensions. The *n*-dimensional SIFT method was proposed in Ref. [36] for extracting and matching salient features from images of arbitrary dimensionality. Volumetric SIFT [37] is an extension of SIFT to extract local volumetric features from 3D voxel model.

Some other adaptations of SIFT for 3D meshes have been published in recent years [38−40]. The method presented in this work is also an extension of SIFT for 3D meshes. The proposed approach will be divided into three main parts in the following subsections: keypoint detection, local feature description, and feature matching. Due to the fact that the proposed approach uses different method to extract local feature, it does not need to assign a canonical orientation, which is necessary in the traditional SIFT method. The 3D meshes $M=(V, T)$ in this work are generated from uniformly sampled point clouds by triangle meshing, where $V=\{\boldsymbol{v}_i\}_{1\le i\le N}$ is the set of mesh vertices and $T=\{t_i\}$ is the set of triangles formed by adjacent vertices.

3.1 Keypoint detection

The keypoint detection is to identify location and scale of these points which are scale space extrema (maxima and minima) and can be repeatable under various cases of the same object. As the SIFT method for 2D images, a scale space in the proposed method is constructed by 3D meshes. The original SIFT method has many octaves, but the proposed method just has one octave. Additionally, a mean filter is utilized rather than a Gaussian filter to build the scale space, because the mean filter is simple relative to the Gaussian filter but they obtain the same performance for face recognition.

The mesh M_{s+1} is obtained by applying a local mean filter to the last scale mesh M_s. Each vertex $\boldsymbol{v}_i^{s+1}$ at the s+1 scale mesh M_{s+1} is computed as follows:

$$\boldsymbol{v}_i^{s+1}=\frac{1}{|N_i|+1}(\boldsymbol{v}_i^s+\sum_{j\in N_i}\boldsymbol{v}_j^s) \tag{1}$$

where N_i represents the set of the first-order neighbors of $\boldsymbol{v}_i^s$ in the three-dimensional space.

A quantitative measure of the shape of a surface at one point p, called the shape index (SI) [41], f_{SI}, is defined as

$$f_{\mathrm{SI}}(p)=\frac{1}{2}-\frac{1}{\pi}\tan^{-1}\frac{k_1(p)+k_2(p)}{k_1(p)-k_2(p)} \tag{2}$$

where k_1 and k_2 are the principal curvatures of the surface, with $k_1 \geq k_2$; f_{SI} describes the attribute of local surface and is insensitive to pose changes. To detect salient points in the scale space, the shape index is computed for each vertex i at each scale s in the scale space. Differences between subsequent scales are given by

$$\mathrm{d}f_{\mathrm{SI}_i}^s=f_{\mathrm{SI}_i}^{s+1}-f_{\mathrm{SI}_i}^s \tag{3}$$

In the difference scale space, each point is selected as feature keypoint if it has higher or lower value of $\mathrm{d}f_{\mathrm{SI}_i}^s$ than all of its neighbours on the same scale as well as on the upper and lower scale, as shown in Fig. 1. The red point in Fig. 1 represents the current point being processed.

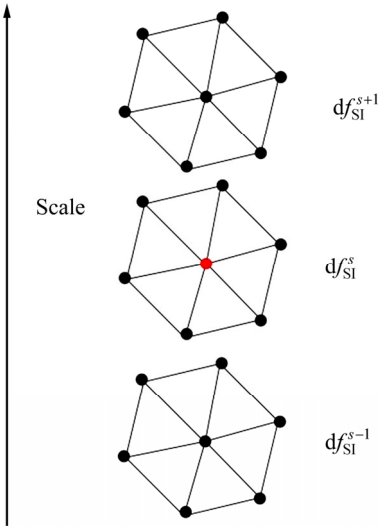

Fig. 1 Neighbourhood of one point in difference scale space

However, there are some keypoints which are volatile and easily affected by variations. For stability, the algorithm of LOWE [32] eliminates the keypoints with low contrast or a strong response along edges for 2D images. The unstable keypoints are also removed before feature extraction in the proposed method. First, a proper sphere with radius R is constructed around each keypoint. When the number of points inside the sphere is lower than one threshold N, the keypoint will be removed. Second, the three eigenvectors $\{\xi_1, \xi_2, \xi_3\}$ extracted using PCA at each keypoint are calculated, which are corresponding to the three eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$ from large to small order. In Section 3.2, a local coordinate system will be established for each keypoint based on the three eigenvectors, as shown in Fig. 2. In the case of $\lambda_1 \approx \lambda_2$ or $\lambda_2 \approx \lambda_3$, noise might cause $\{\xi_1, \xi_2, \xi_3\}$ to swap places. In order to avoid the exchange of coordinate axis, it is required that

$$\lambda_1 >> \lambda_2 >> \lambda_3 \tag{4}$$

$\lambda_1 \geq 5\lambda_2$ and $\lambda_2 \geq 5\lambda_3$ are set empirically in this work.

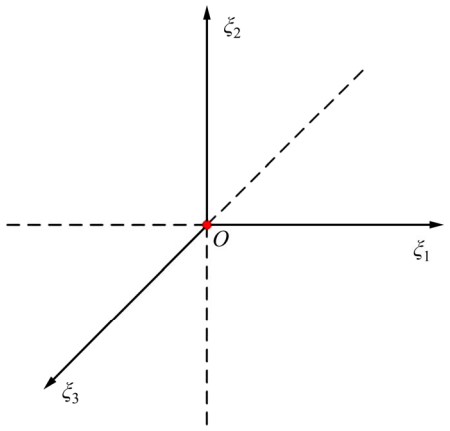

Fig. 2 Local coordinate system at one keypoint

3.2 Local feature description

After obtaining the keypoints with location and scale, a local three-dimensional coordinate system is established using the three eigenvectors $\{\xi_1, \xi_2, \xi_3\}$ for each keypoint, which the keypoint is the origin point. The coordinate system is divided into four parts based on the coordinate value at the direction of ξ_1 and ξ_2 coordinate axes. A sphere with radius $r=(\sqrt{2}/4)R$ is constructed in each part and the distance from its center to the origin is r=0.5R. Meanwhile, a sphere with radius r_o=0.5R is made around the origin. Figure 3 illustrates the top plan view of location of each sphere in the local coordinate system. The number in Fig. 3 represents the order of each sphere in the three-dimensional coordinate system.

In this work, two local and simple geometric features are developed: one distance d and one angle α, as shown in Fig. 4. When the random point p is inside one of spheres with the order from the first to fourth, O_1 is the center of the corresponding sphere, as shown in Fig. 4(a). Otherwise, O_1 is the point (0, 0, 0.5R) when the random point p is inside the sphere around the origin O (the current keypoint), as shown in Fig. 4(b). d is the distance between the origin O and the point p inside one of the five spheres. α is the angle between the line determined by the origin O and the point p and the line determined by the origin O and the point O_1. Because

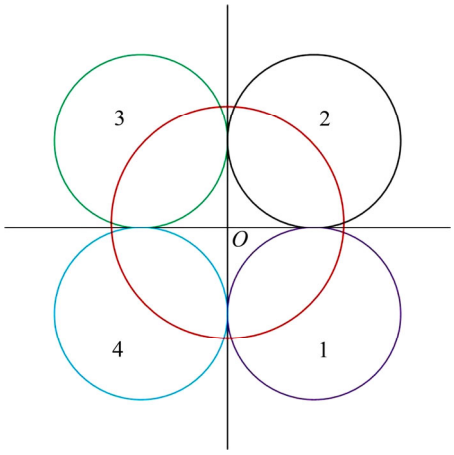

Fig. 3 Top plan view of location of each sphere in local coordinate system

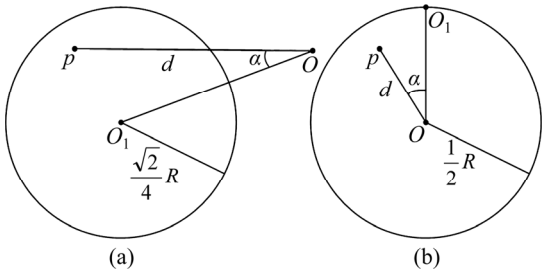

Fig. 4 Distance feature and angle feature for one keypoint: (a) Random point p is inside one of spheres with order from the first to fourth; (b) Random point p is inside sphere around origin

these two features cannot match between faces directly, they are quantized into histograms. Two histograms $\boldsymbol{h}_d$ and $\boldsymbol{h}_\alpha$ are calculated in each sphere. Finally, the histograms are concatenated in a single-feature vector $\boldsymbol{f}=[\boldsymbol{h}_{d0}\ \boldsymbol{h}_{\alpha0}\ \cdots\boldsymbol{h}_{d4}\ \boldsymbol{h}_{\alpha4}]$. The feature vector $\boldsymbol{f}$ describes the local properties of the corresponding keypoint.

Self-occlusion caused by large pose changes and external partial occlusion can both result in loss of information. To reduce the influence by these problems, the approximate left-right symmetry in human face is utilized for increasing amount of information. It can improve the integrity and robustness of facial information in these cases. The feature vector is symmetrized by flipping these spheres at the left-right direction. Figure 5 illustrates the top plan view of locations of the five spheres after symmetrization. It can be seen that only the location of the sphere around the origin is unchanged. Therefore, the symmetrical feature $\boldsymbol{f}^{\mathrm{s}}$ corresponding to the feature $\boldsymbol{f}$ can be obtained according to Fig. 5.

$$\boldsymbol{f}^{\mathrm{s}}=[\boldsymbol{h}_{d0}\ \boldsymbol{h}_{\alpha0}\ \boldsymbol{h}_{d4}\ \boldsymbol{h}_{\alpha4}\ \boldsymbol{h}_{d3}\ \boldsymbol{h}_{\alpha3}\ \boldsymbol{h}_{d2}\ \boldsymbol{h}_{\alpha2}\ \boldsymbol{h}_{d1}\ \boldsymbol{h}_{\alpha1}] \quad (5)$$

The symmetrical features are finally added to the feature descriptors of the original face and the feature set is extended.

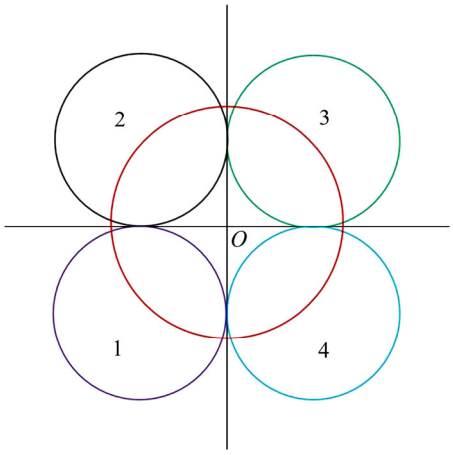

Fig. 5 Top plan view of locations of five spheres after symmetrization

3.3 Feature matching

To find corresponding keypoints between two facial surfaces, Euclidean distance is used as similarity measure to compare the histogram feature vectors of both surfaces. For the feature vector of one keypoint in one surface, the distances with all feature vectors of other surfaces are computed and ranked in ascending order. A match is accepted if the ratio between the first and the second is smaller than one threshold r_{th}; otherwise, the match will be rejected. Matches are mostly found between two facial surfaces of the same subject. The number of matches is used as similarity criterion.

4 Experiments

The main purpose of the proposed approach is to solve 3D face recognition problem under expression variations, external partial occlusions and pose changes. To demonstrate the effectiveness of the proposed method, three public datasets are used for experiments: the Bosphorus database [42], the BU-3DFE database [43], and the Gavab database [44].

4.1 Parameter setting

Before presenting test results on the three 3D face datasets, some parameters need to be determined, including R (mentioned in Section 3.1), N (mentioned in Section 3.1), and r_{th} (mentioned in Section 3.3). R, N, and r_{th} are chosen from three sets {6, 7, 8, 9, 10}, {10, 15, 20, 25, 30}, and {0.4, 0.5, 0.6, 0.7, 0.8}, respectively. To obtain good parameters, the facial scans of the first 35 objects in the Bosphorus database are adopted for experiments. After doing experiments in every possible combination of the three parameters, the best combination is found when R, N, and r_{th} are equal to 8, 20, and 0.6, respectively.

4.2 Experiments on Bosphorus dataset

The Bosphorus database incorporates a rich set of expressions, various kinds of poses, and different types of occlusions which may occur in real life. The facial data are acquired with a 3D digitizer device based on a commercial structured-light. The sensor resolutions in x, y and z dimensions are 0.3, 0.3 and 0.4 mm, respectively. The database consists of 4666 scans from 105 subjects, 61 men and 44 women. The majority of the subjects are aged between 25 and 35. The Bosphorus database contains two different types of facial expressions. The first type is based on facial action units (AUs) of the facial action coding system [45]. It contains 28 different action units. They are grouped into three sets: 20 lower face AUs, five upper face AUs and three AU combinations. There are six basic emotional expressions in the second type: happiness, surprise, fear, sadness, anger and disgust. For each subject, three types of head poses are collected which contain five yaw rotations, four pitch rotations, and two cross rotations which incorporate both yaw and pitch. In addition, four types of occluded scans are acquired: mouth occlusion by hand, face with eyeglasses, face occluded with hair, occlusion of left eye, and forehead region by hands. Figure 6 shows some examples of one subject in this dataset.

In the Bosphorus 3D database, each vertex on a 3D facial surface is described by 3D spatial coordinates and the corresponding RGB color image. But only 3D spatial scattered coordinates with x, y, and z values are utilized in this work. Before detecting keypoints, some preprocessing steps are applied to the original face point clouds. A median filter is used to remove the spikes and a Gaussian filter is applied to smoothing the facial data. Then, holes are filled by cubic interpolation. For keeping a fixed correspondence across the different face point clouds, a uniform resampling is applied to all face point clouds. A simple triangle-based linear interpolation is used with 1 mm resolution to finish uniform resampling. Furthermore, the depth value is smoothed by a Gaussian filter again.

Experiments are carried out on the expression subset, the occlusion subset, the rotation subset and the whole set of the Bosphorus database, respectively. Table 1 illustrates the rank-1 recognition rates (RR) of these four experiments. The rank-1 RR is the percentage of all probes for which the best match in the gallery is associated to the same object, which is a good measure for identification. In these experiments, the gallery set consists of the first neutral scan of each subject.

Table 1 Rank-1 recognition rates on Bosphorus database

Datum	Rank-1 recognition rate/%
Expression subset	98.1
Occlusion subset	96.8
Pose variations subset	92.3
Whole set	96.3

Table 2 shows the comparison of the rank-1 recognition rates of the proposed approach and some

Fig. 6 Some examples of one subject in Bosphorus database

Table 2 Comparison of rank-1 recognition rates by different methods on Bosphorus database

Approach	Data	Year	Probe	Gallery	Rank-1 recognition rate/%
ALYUZ et al [46]	Expression	2010	2815	105	98.19
HUANG et al [47]	Expression + occlusion	2012	3196	105	97.0
SMEETS et al [40]	All	2012	4561	105	93.7
LIU et al [48]	All	2013	−	47	95.6
Proposed approach	Expression	2014	2815	105	98.1
Proposed approach	Expression + occlusion	2014	3196	105	97.5
Proposed approach	All	2014	4561	105	96.3

other state-of-the-art methods on the Bosphorus database. The rank-1 recognition rate in Ref. [46] has 0.09% higher than that of the proposed approach on the expression subset of the Bosphorus database. However, it does not consider the cases under occlusions and pose changes. For the expression and occlusion subset, the proposed method obtains better performance than the algorithm in Ref. [47]. Moreover, the proposed method achieves higher recognition rate than the algorithm in Refs. [40, 48] on the whole set of the Bosphorus database.

4.3 Experiments on BU-3DFE dataset

The BU-3DFE database is especially used for expression recognition and expression invariant face recognition. It contains 2500 facial scans from 100 persons with ages ranging from 18 to 70 years. The facial shape is acquired using a stereoscopic scanner that captures the face of subject using six digital cameras and merges the camera data to produce a single 3D triangular mesh. Each subject has a neutral expression and six basic emotional expressions. And the six basic emotional expressions have four different levels of intensities. The intensity ranges from low, middle, high, and highest. Figure 7 shows six basic emotional expressions of one subject at intensity level 3 and level 4.

The 3D facial data of the BU-3DFE database have already been preprocessed by the provider. So, the only thing required is to do a uniform resampling on face point cloud data before keypoint detection. Similar to the experiments on the Bosphorus database, the gallery set is constructed by using the first neutral scan of each subject and all the rest ones are made up of the probe set. Therefore, the size of gallery set and probe set are 100 and 2400, respectively.

Table 3 lists the rank-1 recognition rates of different approaches on the BU-3DFE database for comparison. It can be seen from Table 3 that the proposed approach achieves a rank-1 recognition rate up to 98.4% and out performs most of the existing state-of-the-art methods. Although the performance of algorithm in Ref. [49] is slightly superior to that of the proposed approach, the average number of each subject in the gallery set in their experiments is 6.95, which is more than one.

4.4 Experiments on Gavab dataset

Gavab database is one of the most expression-rich and noise-prone 3D face datasets currently available to the public. It contains 61 different subjects and these subjects are all Caucasian with 45 males and 16 females. The 3D scans are acquired by a Minolta Vi-700 laser without control of the light condition. Each subject has 9 scans for different poses and expressions. The scans with pose variations contain one facial scan while looking up (35º), one while looking down (−35º), one for the right profile (90º), one for the left profile (−90º) as well as one with random pose. The expression scans include one with a smile and one with a pronounced laugh. Figure 8 shows some examples of faces in this dataset.

A similar preprocessing step is utilized in the Bosphorus database. Like the above two databases, the gallery set in the experiments on the Gavab database is also made up of one neutral scan of each subject. First, experiments are operated on the neutral and expressive subsets of the Gavab database. Table 4 demonstrates the

(a)

(b)

Fig. 7 Six basic expressions at intensity level 3 (a) and level 4 (b) of same subject in BU-3DFE database

Table 3 Comparison of rank-1 recognition by different methods in BU-3DFE database

Approaches	Year	Probe	Gallery	Number of subjects	Rank-1 recognition rate/%
KAUSHIK et al [49]	2009	695	695	100	98.5
DANIYAL et al [50]	2009	−	>1 for each subject	−	96.5
WANG et al [51]	2010	−	>1 for each subject	−	98
SMEETS et al [52]	2010	900	100	100	94.5
LEI et al [53]	2013	−	> 1 for each subject	100	97.7
Proposed approach	2014	2400	100	100	98.4

Fig. 8 Examples of all 3D scans of same subject from GAVAB dataset [54]

comparison with different methods on this subset. It can be seen that the proposed approach surpasses all other methods for the neutral and expressive faces. The rank-1 recognition rate of the proposed approach reaches up to 96.43%. Second, experiments are carried out on the subset with pose variations. The comparison with different methods on this subset is shown in Table 5. Table 5 illustrates that the proposed approach has better performance than other methods.

Table 4 Comparison of rank-1 recognition rates by different methods on neutral and expressive subsets of Gavab database

Approach	Year	Rank-1 recognition rate/%
LI et al [54]	2009	94.68
DRIRA et al [55]	2010	94.67
HUANG et al [47]	2012	95.49
Proposed approach	2014	96.43

Table 5 Comparison of rank-1 recognition rates with different methods on subset with pose variations of Gavab database

Approach	Year	Rank-1 recognition rate/%
DRIRA et al [55]	2010	88.94
HUANG et al [47]	2012	91.39
Proposed approach	2014	91.88

4.5 Summary

The proposed approach is validated on three public datasets, i.e. the Bosphorus, the BU-3DFE and the Gavab databases. The results obtained in these experiments show that the proposed approach is comparable to the best method so far presented in the literature. Moreover, the proposed approach can deal with 3D face recognition under rich facial expressions, external partial occlusions and large pose variations.

5 Conclusions

A novel approach is presented to cope with 3D face recognition problem under varying expressions, external occlusions and pose variations. SIFT features are extracted on 3D meshes and compare the features between different facial surfaces for matching faces. A difference scale space is built for detecting salient points. Shape index extrema are selected as keypoints in this scale space. Then, two steps are utilized to eliminate the volatile keypoints. After determining the keypoints, a local coordinate system for each keypoint is established based on the eigenvectors obtained by PCA. Two local geometric features are computed around each keypoint on the basis of the local coordinate system and they are good representations for 3D facial scans. To mitigate the loss of information caused by large pose change and external partial occlusion, the features are expanded by symmetrization based on the left-right symmetry in human face.

The proposed method is evaluated on the Bosphorus, BU-3DFE, and Gavab databases. The proposed algorithm achieves a rank-1 recognition rate of 96.3% on the whole set of Bosphorus database and 98.4% on the BU-3DFE database. Meanwhile, the proposed approach obtains a rank-1 recognition rate of 96.43% and 91.88% on the subset without pose variations and the subset with pose variation of the Gavab database. These results show that the proposed approach can be comparable to some existing state-of-the-art algorithms and even outperforms some of them.

References

[1] JAIN A K, ROSS A, PRABHAKAR S. An introduction to biometric recognition [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14(1): 4−20.

[2] JAIN A K, ROSS A, PRABHAKAR S. Biometric identification [J]. Communications of the ACM, 2000, 43(2): 91−98.

[3] MEDIONI G, WAUPOTITSCH R. Face modeling and recognition in 3-D [C]// Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures. Nice, France: IEEE, 2003: 232−233.

[4] CHANG K, BOWYER K, FLYNN P. A survey of approaches and challenges in 3D and multi-modal 2D+3D face recognition [J]. Computer Vision and Image Understanding, 2006, 101(1): 1−15.

[5] BESL P J, MCKAY N D. A method for registration of 3-D shapes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(2): 239−256.

[6] LU Xiao-guang, JAIN A K, COLBRY D. Matching 2.5D face scans to 3D models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(1): 31−43.

[7] MIAN A S, BENNAMOUN M, OWENS R. An efficient multimodal 2D-3D hybrid approach to automatic face recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(11): 1927−1943.

[8] MAHOORA M, ABDEL-MOTTALEB M. Face recognition based on 3D ridge images obtained from range data [J]. Pattern Recognition, 2009, 42(3): 445−451.

[9] CHANG K I, BOWYER K W, FLYNN P J. Multiple nose region matching for 3D face recognition under varying facial expression [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(10): 1695−1700.

[10] WANG Yue-ming, PAN Gang, WU Zhao-hui, WANG Yi-gang. Exploring facial expression effects in 3D face recognition using partial ICP [C]// Proceedings of the 7th Asian Conference on Computer Vision. Berlin, Heidelberg: Springer, 2006: 581−590.

[11] SALAH A A, ALYUZ N, AKARUN L. Registration of three-dimensional face scans with average face models [J]. Journal of Electronic Imaging, 2008, 17(1): 1−14.

[12] HUTTENLOCHER D P, KLANDERMAN G A, RUCKLIDGE W J. Comparing images using the Hausdorff distance [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(9): 850−863.

[13] ACHERMANN B, BUNKE H. Classifying range images of human faces with Hausdorff distance [C]// Proceedings of the 15th International Conference on Pattern Recognition. Barcelona, Spain: IEEE, 2000: 809−813.

[14] LEE Y, SHIM J. Curvature based human face recognition using depth weighted Hausdorff distance [C]// Proceedings of the 2004 International Conference on Image Processing. Singapore: IEEE, 2004: 1429−1432.

[15] RUSS T D, KOCH M W, LITTLE C Q. A 2D range Hausdorff approach for 3D face recognition [C]// Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005: 1−8.

[16] CONDE C, RODRIGUEZ-ARAGON L J, CABELLO E. Automatic 3D face feature points extraction with spin images [C]// Proceedings of the Third International Conference on Image Analysis and Recognition. Povoa de Varzim, Portugal, 2006: 317−328.

[17] ROMERO M, PEARS N. Point-pair descriptors for 3d facial landmark localization [C]// Proceedings of the IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems. Washington, DC, USA: IEEE, 2009: 1−6.

[18] PERAKIS P, PASSALIS G, THEOHARIS T, KAKADIARIS I A. 3D facial landmark detection under large yaw and expression variations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(7): 1552−1564.

[19] CHUA C, JARVIS R. Point signatures: A new representation for 3D object recognition [J]. International Journal of Computer Vision, 1997, 25(1): 63−85.

[20] WANG Ying-jie, CHUA Chin-seng, HO Yeong-khing. Facial feature detection and face recognition from 2D and 3D images [J]. Pattern Recognition Letters, 2002, 23(10): 1191−1202.

[21] IRFANOGLU M O, GOKBERK B, AKARUN L. 3D shape-based face recognition using automatically registered facial surfaces [C]// Proceedings of the 17th International Conference on Pattern Recognition. Cambridge: IEEE, 2004: 183−186.

[22] WANG Ying-jie, CHUA Chin-seng. Robust face recognition from 2D and 3D images using structural Hausdorff distance [J]. Image and Vision Computing, 2006, 24(2): 176−185.

[23] NAGAMINE T, UEMURA T, MASUDA I. 3D facial image analysis for human identification [C]// Proceedings of the 11th IAPR International Conference on Pattern Recognition. The Hague, Netherlands: IEEE, 1992: 324−327.

[24] BEUMIER C, ACHEROY M. Automatic 3D face authentication [J]. Image and Vision Computing, 2000, 18(4): 315−321.

[25] SAMIR C, SRIVASTAVA A, DAOUDI M. Three-dimensional face recognition using shapes of facial curves [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(11): 1858−1863.

[26] COLOMBO A, CUSANO C, SCHETTINI R. 3D face detection using curvature analysis [J]. Pattern Recognition, 2006, 39(3): 444−455.

[27] GOKBERK B, IRFANOGLU M O, AKARUN L. 3D shape-based face representation and feature extraction for face recognition [J]. Image and Vision Computing, 2006, 24(8): 857−869.

[28] GOKBERK B, DUTAGACI H, ULAS A, AKARUN L, SANKUR B. Representation plurality and fusion for 3-D face recognition [J]. IEEE Transactions on Systems, Man, and Cybernetics, 2008, 38(1): 155−173.

[29] TANAKA H T, IKEDA M, CHIAKI H. Curvature-based face surface recognition using spherical correlation—Principal directions for curved object recognition [C]// Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition. Nara, Japan: IEEE, 1998: 372−377.

[30] BRONSTEIN M A, BRONSTEIN M M, KIMMEL R. Three dimensional face recognition [J]. International Journal of Computer Vision. 2005, 64(1): 5−30.

[31] LOWE D G. Object recognition from local scale-invariant features [C]// Proceedings of the Seventh IEEE International Conference on Computer Vision. Kerkyra, 1999: 1150−1157.

[32] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91−110.

[33] SE S, LOWE D G, LITTLE J J. Vision-based mobile robot localization and mapping using scale-invariant features [C]// Proceedings of International Conference on Robotics and Automation. IEEE, 2001: 2051−2058.

[34] BROWN M, LOWE D G. Automatic panoramic image stitching using invariant features [J]. International Journal of Computer Vision, 2007, 74(1): 59−73.

[35] BASHARAT A, ZHAI Y, SHAH M. Content based video matching using spatiotemporal volumes [J]. Computer Vision and Image

Understanding, 2008, 110(3): 360−377.

[36] CHEUNG W, HAMARNEH G. *n*-SIFT: *n*-dimensional scale invariant feature transform [J]. IEEE Transactions on Image Processing, 2009, 18(9): 2012−2021.

[37] OSADA K, FURUYA T, OHBUCHI R. Shrec'08 entry: Local volumetric features for 3D model retrieval [C]// Proceedings of IEEE International Conference on Shape Modeling and Applications. New York, USA: IEEE, 2008: 245−246.

[38] ZOU Guang-yu, HUA Jing, DONG Ming, QIN Hong. Surface matching with salient keypoints in geodesic scale space [J]. Computer Animation and Virtual Worlds, 2008, 19(3/4): 399−410.

[39] ZAHARESCU A, BOYER E, VARANASI K, HORAUD R. Surface feature detection and description with applications to mesh matching [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Miami, Florida: IEEE, 2009: 373−380.

[40] SMEETS D, KEUSTERMANS J, VANDERMEULEN D, SUETENS P. meshSIFT: Local surface features for 3D face recognition under expression variations and partial data [J]. Computer Vision and Image Understanding, 2013, 117(2): 158−169.

[41] DORAI C, JAIN A K. COSMOS—A representation scheme for 3D free-form objects [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 19(10): 1115−1130.

[42] SAVRAN A, ALYUZ N, DIBEKLIOGLU H, CELIKTUTAN O, GOKBERK B, SANKUR B, AKARUN L. Bosphorus database for 3D face analysis [J]. Biometrics and Identity Management, 2008, 5372: 47−56.

[43] YIN Li-jun, WEI Xiao-zhou, SUN Yi, WANG Jun, ROSATO M J. A 3D facial expression database for facial behavior research [C]// Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition. Southampton, 2006: 211−216.

[44] MORENO A B, SANCHEZ A. GavabDB: A 3D face database [C]// Proceedings of COST Workshop on Biometrics on the Internet: Fundamentals, Advances and Applications. Vigo, Spain, 2004: 77−82.

[45] KMAN P, FRIESEN W V. The facial action coding system: A technique for the measurement of facial movement [M]. San Francisco: Consulting Psychologists Press, 1978.

[46] ALYUZ N, GOKBERK B, AKARUN L. Regional registration for expression resistant 3-D face recognition [J]. IEEE Transactions on Information Forensics and Security, 2010, 5(3): 425−440.

[47] HUANG Di, ARDABILIAN M, WANG Yun-hong, CHEN Li-ming. 3-D face recognition using eLBP-based facial description and local feature hybrid matching [J]. IEEE Transactions on Information Forensics and Security, 2012, 7(5): 1551−1565.

[48] LIU Pei-jiang, WANG Yun-hong, HUANG Di, ZHANG Zhao-xiang, CHEN Li-ming. Learning the spherical harmonic features for 3-D face recognition [J]. IEEE Transactions on Image Processing, 2013, 22(3): 914−925.

[49] KAUSHIK V D, BUDHWAR A, DUBEY A, AGRAWAL R, GUPTA S, PATHAK V K, GUPTA P. An efficient 3D face recognition algorithm [C]// Proceedings of the 3rd International Conference on New Technologies, Mobility and Security. Cairo, 2009: 1−5.

[50] DANIYAL F, NAIR P, CAVALLARO A. Compact signatures for 3D face recognition under varying expressions [C]// Proceedings of the Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance. Genova, 2009: 302−307.

[51] WANG Yue-ming, LIU Jian-zhuang, TANG Xiao-ou. Robust 3D face recognition by local shape difference boosting [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(10): 1858−1870.

[52] SMEETS D, FABRY T, HERMANS J, VANDERMEULEN D, SUETENS P. Fusion of an isometric deformation modeling approach using spectral decomposition and a region-based approach using ICP for expression-invariant 3D face recognition [C]// Proceedings of the 20th International Conference on Pattern Recognition. Istanbul, 2010: 1172−1175.

[53] LEI Yin-jie, BENNAMOUN M, EL-SALLAM A A. An efficient 3D face recognition approach based on the fusion of novel local low-level features [J]. Patter Recognition, 2013, 46(1): 24−37.

[54] DRIRA H, AMOR B B, DAOUDI M, SRIVASTAVA A. Pose and expression-invariant 3D face recognition using elastic radial curves [C]// Proceedings of British Machine Vision Conference. Aberystwyth, UK, 2010: 1−11.

[55] LI Xiao-xing, JIA Tao, ZHANG Hao. Expression-insensitive 3D face recognition using sparse representation [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Miami, Florida, 2009: 2575−2582.

(Edited by YANG Hua)