

RESEARCH

Open Access

Three-dimensional face recognition under expression variation

Xueqiao Wang*, Qiuqi Ruan, Yi Jin and Gaoyun An

Abstract

In this paper, we introduce a fully automatic framework for 3D face recognition under expression variation. For 3D data preprocessing, an improved nose detection method is presented. The small pose is corrected at the same time. A new facial expression processing method which is based on sparse representation is proposed subsequently. As a result, this framework enhances the recognition rate because facial expression is the biggest obstacle for 3D face recognition. Then, the facial representation, which is based on the dual-tree complex wavelet transform (DT-CWT), is extracted from depth images. It contains the facial information and six subregions' information. Recognition is achieved by linear discriminant analysis (LDA) and nearest neighbor classifier. We have performed different experiments on the Face Recognition Grand Challenge database and Bosphorus database. It achieves the verification rate of 98.86% on the all vs. all experiment at 0.1% false acceptance rate (FAR) in the Face Recognition Grand Challenge (FRGC) and 95.03% verification rate on nearly frontal faces with expression changes and occlusions in the Bosphorus database.

Keywords: Dual-tree complex wavelet transform; 3D face recognition; Sparse representation; Linear discriminant analysis

1 Introduction

3D face recognition is a continuously developing subject with many challenging issues [1-3]. These years, many new 3D face recognition methods which were demonstrated on the Face Recognition Grand Challenge (FRGC) v2 data have got good performances.

Regional matching scheme was firstly proposed by Faltemier et al. [4]. In their paper, the whole 3D face images were divided into 28 patches. The fusion results from independently matched regions could achieve good performance. Wang et al. [5] extracted the Gabor, LBP, and Haar features from the depth image, and then the most discriminative local feature was selected optimally by boosting and trained as weak classifiers for assembling three collective strong classifiers. Mian et al. [6] extracted the spherical face representation (SFR) of the 3D facial data and the scale invariant feature transform (SIFT) descriptor of the 2D data to train a rejection classifier. The remaining faces were verified using a region-based matching approach which was robust to facial expression. Berretti et al. [7] proposed an approach that took into

account the graph form to reflect geometrical information for 3D facial surface, and the relevant information among the neighboring points could be encoded into a compact representation. 3D weighted walkthrough (3DWW) descriptors were proposed to demonstrate the mutual spatial displacement among pairwise arcs of points of the corresponding stripes. Zhang et al. [8] found a novel resolution invariant local feature for 3D face recognition. Six different scale invariant similarity measures were fused at the score level, which increased the robustness against expression variation.

The accuracy of 3D face recognition could be significantly degraded by large facial expression variations. Alyuz et al. [9] proposed an expression resistant 3D face recognition method based on the regional registration. In recent years, many methods dealt with facial expression before recognition. Kakadiaris et al. [10] utilized the elastically adapted deformable model firstly, and then they mapped the 3D geometry information onto a 2D regular grid, thus combining the descriptiveness of the 3D data with the computational efficiency of the 2D data. A multistage fully automatic alignment algorithm and the advanced wavelet analysis were used for recognition. Drira et al. [11] represented facial surfaces by radial curves emanating from the nose tips and used elastic shape

* Correspondence: silvia.wxq@gmail.com

Beijing Key Laboratory of Advanced Information Science and Network Technology, Institution of Information Science, Beijing Jiaotong University, Beijing 100044, China

analysis of these curves to develop a Riemannian framework for analyzing shapes of full facial surfaces. Their method used the nose tips which are already provided. Mohammadzade et al. [12] presented a new iterative method which can deal with 3D faces with opened mouth. They performed experiments to prove that the combination of the normal vectors and the point coordinates can improve the recognition performance. A verification rate of 99.6% at a false acceptance rate (FAR) of 0.1% has been achieved using the proposed method for the all versus all experiment. Amberg et al. [13] described an expression invariant method for face recognition by fitting an identity/expression separated 3D Morphable Model to shape data. The expression model greatly improved recognition. Their method operated at approximately 40 to 90 s per query.

Our method is an automatic method for 3D face recognition. The framework of our method is presented in Figure 1. For data preprocessing, an improved nose detection method is proposed. At the same time, the small pose of the face can be corrected. Then, the face region (face without hair and ears) is gotten using a sphere centered at the nose tip. After finding the face region, the facial expression is removed using a new method which is based on sparse representation. Finally, the depth image is constructed. In the training section, we

use all the 943 faces in FRGC v1 for training. First of all, we extract the four-level magnitude subimages of each training faces using DT-CWT. Subsequently, we vectorize the six magnitude subimages into a large vector which dimension is 384 and utilize the linear discriminant analysis (LDA) [14] to learn the subspace of the training faces and then record the transformation matrix. Secondly, the six subregions' four-level magnitude subimages are extracted using DT-CWT, and they are vectorized into a large vector which dimension is 2,304. After that, we utilize the linear discriminant analysis [14] to learn the transformation matrix too. Finally, we get all the gallery faces' two features using DT-CWT and their transformation matrix, respectively, to establish two LDA subspaces. In the testing section, we obtain all the probe faces' two features by using DT-CWT and their two transformation matrices, respectively. Cosine distance is used to establish two similarity matrices. In the end of the method, two similarity matrices are fused, and the nearest neighbor classifier is used to finish the recognition process.

The main contributions of this work can be summarized as follows:

- The first contribution is an improved nose detection method which can correct the small pose of the face iteratively. The proposed nose detection algorithm is

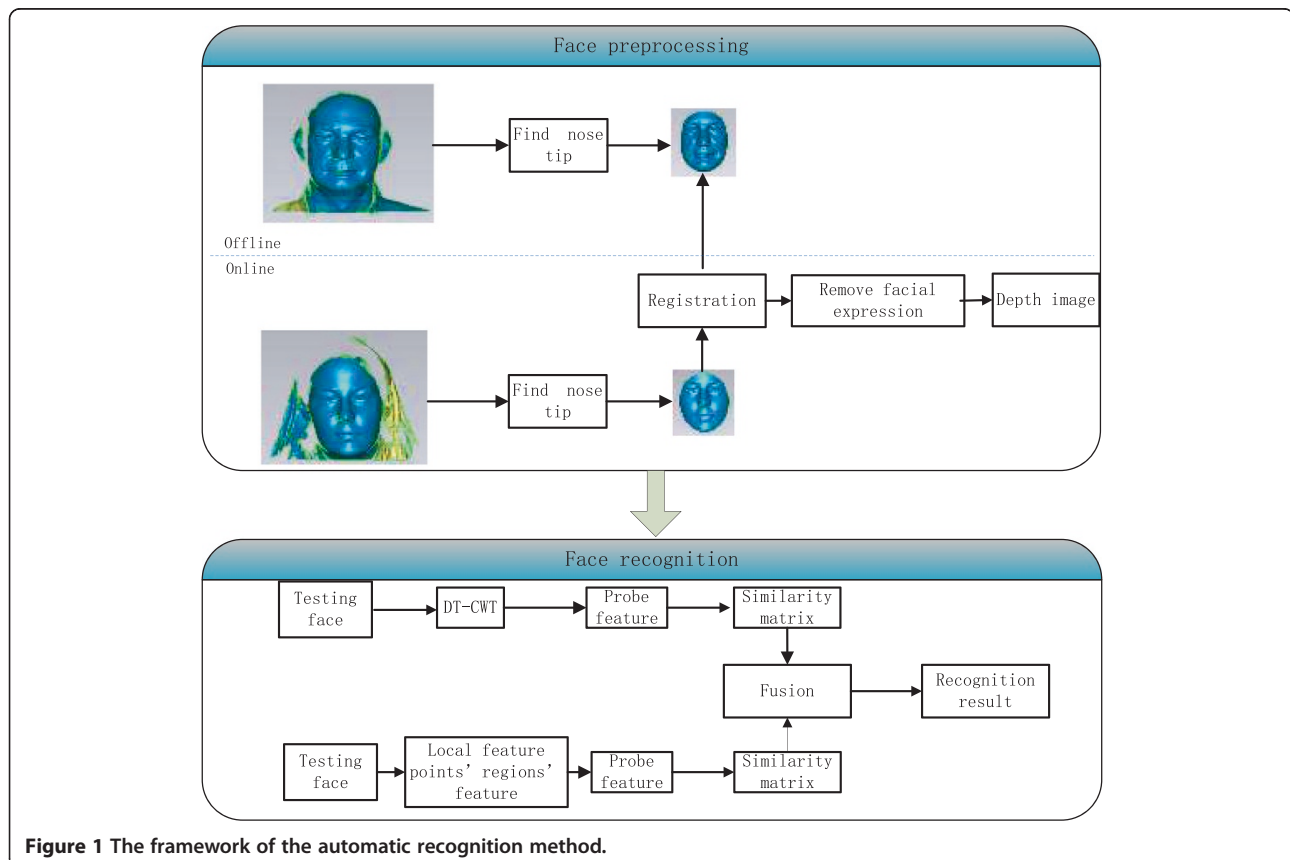


Figure 1 The framework of the automatic recognition method.

simple, and the success rate is 99.95% in the FRGC database.

- The second one is that we propose a new 3D facial expression processing method which is based on sparse representation. Li et al. [15] utilized sparse representation into 3D face recognition, but they applied it in the recognition section. In this paper, sparse representation is used for facial expression processing. The objective of the sparse representation is to relate a probe with the minimum number of gallery dataset. Considering that the first task of our expression processing work is to find the minimum number of expressional components out of the dictionary (because people only make one expression for one time), the objective of sparse representation is naturally better suited for finding the expressional deformation from the dataset. This method is a learning method that can abstract the testing face's neutral component from a dictionary of neutral and expressional spaces, and it only costs 14.91 s for removing one facial expression (The type of our CPU is Intel (R) Core (TM) i3-2120, and the RAM is 2 GB.). The proposed method is more simple and only cost less time.

The paper is organized as follows: In Section 2, the data preprocessing methods are proposed. The improved nose tip detection method is presented in this section. Then, the 3D facial expression processing method is presented in Section 3. In Section 4, the framework of our 3D face recognition method is given. Experimental results are given in Section 5, and the conclusions are drawn in Section 6.

2 3D data preprocessing

Firstly, a 3×3 Gaussian filter is used to remove spikes and noise, and then the range data are subsampled at a 1:4 ratio.

Some 3D faces in the FRGC database contains information of the ears, while some faces' ears are hidden by the hair. For the purpose of consistency, we only use the face region into recognition. Now, we introduce the face region extracting method.

2.1 Nose detection

The nose is the center of a 3D face, so nose detection is important for facial region extraction. The block diagram of the proposed procedure for nose detection is presented in Figure 2.

In this paper, the first step of nose tip detection is finding the central stripe. Details are presented in our earlier work [16].

We use the face with ID 02463d453 in FRGC v1 as the standard face and manually find its nose tip on its stripe. Subsequently, we find other persons' nose tip using an automatic iterative algorithm. Let us suppose that A is

the central stripe of the face with ID 02463d453, and B is the central stripe of the face whose nose tip needs to be found. The method is as follows:

- (1) Align stripe A to stripe B using the ICP [17] method and record the transformation matrix M_2 .
- (2) Use M_2 to find point p which is the first person's transformed nose tip.
- (3) Crop a sphere (radius = 37 mm) centered at point p . The highest point in the sphere is found as the nose tip of B. The step is shown in our previous work [16].
- (4) Crop a sphere (radius = 90 mm) centered at the nose tip and align to the standard face. Calculate the transformed nose tip p_1 .
- (5) Crop a sphere (radius = 25 mm) centered at point p_1 . The highest point in the sphere is found as the new nose tip p_2 .
- (6) If $\|p_2 - p_1\| < 2$ mm, p_2 is the nose tip, else, back to step (4).

2.2 Face region

Once the nose tip is successfully found, the region in the last step of nose detection is used as the face region. All the faces with excessive head rotation, hair artifact, and big expressions were successfully segmented by the proposed nose detection algorithm. Some examples are presented in Figure 2.

3 3D facial expression processing method based on sparse representation

Facial expression is one of the biggest obstacle of 3D face recognition because 3D face has less information and some information on the face can be changed easily by facial expression. In this section, we introduce a new expression processing method for removing facial expression which is based on sparse representation. We expect that our method could establish correspondence between an open mouth and estimated neutral component.

3.1 Brief introduction of sparse representation

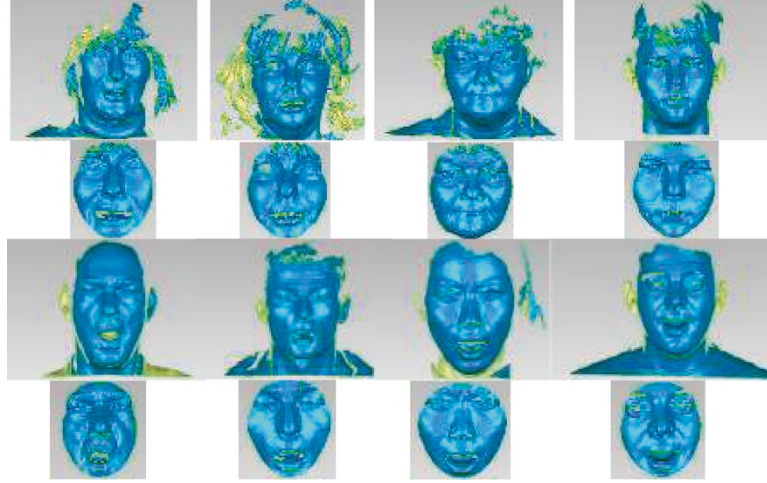
In this paper, we use L1-regularized least squares regression [6,18] to estimate the coefficients of our model. L1-regularized is known to produce sparse coefficients and can be robust to irrelevant features. In general, the problem can be formulated as:

$$\hat{a} = \operatorname{argmin} \left\{ \|y - Ax\|_2^2 + \gamma \|x\|_1 \right\} \quad (1)$$

where y is the test sample, x is the sparse representation on dictionary A , and γ is a scalar constant (we use $\gamma = 5,000$ in this paper). The feature-sign search method [6] is adopted to solve Equation 1.



(A) Example of the face with excessive head rotation. After three times nose verification, the pose of the face is corrected.



(B) Examples of the faces with hair artifact and big expression.

Figure 2 All such faces were successfully segmented by the proposed nose detection algorithm.

3.2 Facial expression processing method

First of all, we use a triangle-based linear interpolation method to fit a surface $Z=f(X, Y)$ (the size of it is 128×128). Meanwhile, we use a triangle-based linear interpolation to fit a surface too (the size of it is 384×384), and then we establish the depth image using the surface for the feature extraction in Section 4.

We consider face ($F_{face}=Z$) as the sum of a reference face ($F_{reference}$), a neutral component ($\Delta F_{Neutral}$), and an expressional component ($\Delta F_{Expression}$). In this paper, we use the face with ID F0001_NE00 of BU-3DFE dataset as reference.

$$F_{face} = F_{reference} + \Delta F_{Neutral} + \Delta F_{Expression} \quad (2)$$

The goal of this section is getting $F_{Neutral}$:

$$F_{Neutral} = F_{reference} + \Delta F_{Neutral} \quad (3)$$

In this paper, we use sparse representation to evaluate the testing face's $\Delta F_{Neutral}$ and $\Delta F_{Expression}$ from a neutral space and an expressional space, respectively, because we want to find the minimum number of expressional components out of the dictionary and a linear combination of $\Delta F_{Neutral}$ using a neutral space. First of all, the dictionary $A = [A_1, A_2]$ is needed to be established, where A_1 is a neutral space and A_2 is an expressional space.

The results of 275 neutral faces (each person's first face of FRGC v1) subtracting the reference face, respectively, and then vectorizing into 275 large vectors are used to be the neutral space $A_1 = [A_1^1, A_1^2, \dots, A_1^{275}]$, where

$$A_1^i = F_{neutral}^i - F_{reference} = \begin{pmatrix} \Delta z_1^1 \\ \Delta z_1^2 \\ \vdots \\ \Delta z_1^n \end{pmatrix}. \text{ Then, the re-}$$

sults of 460 different expressional faces (the first 10 men's 23 expressional faces from BU-3DFE dataset and the first 10 women's 23 expressional faces from BU-3DFE dataset) subtracting their corresponding neutral face, respectively, and vectorizing into 460 large vectors are applied to be the expressional space $A_2 = [A_2^1, A_2^2, \dots, A_2^{460}]$, where

$$A_2^i = F_{expression}^i - F_{neutral} = \begin{pmatrix} \Delta z_2^1 \\ \Delta z_2^2 \\ \vdots \\ \Delta z_2^n \end{pmatrix}.$$

The reference face and the first persons' expressional faces are shown in Figure 3.

In the testing section, the reference face is subtracted from the testing face and the result's sparse representation of the dictionary A is abstracted by Equation 1 $\hat{x} = \arg$

$$\min \left\{ \left\| y - [A_1, A_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_2^2 + \gamma \left\| x_1 \right\|_1 \right\}, \text{ where}$$

$$y = F_{test} - F_{reference} = \begin{pmatrix} \Delta z_{test}^1 \\ \Delta z_{test}^2 \\ \vdots \\ \Delta z_{test}^n \end{pmatrix}. \text{ Because neutral}$$

components of neutral faces are highly correlated, this method can find the familiar neutral components of the testing face. After this, we reconstruct the testing face's neutral component using A_1 and sparse vector \hat{x}_1 :

$$\Delta \hat{F}_{Neutral} = A_1 \hat{x}_1 \quad (4)$$

So, $\hat{F}_{Neutral}$ is equal to the sum of $F_{reference}$ and $A_1 \hat{x}_1$:

$$\begin{aligned} \hat{F}_{Neutral} &= F_{reference} + \Delta \hat{F}_{Neutral} \\ &= F_{reference} + A_1 \hat{x}_1 \end{aligned} \quad (5)$$

But $\hat{F}_{Neutral}$ is approximate, so each point in $\hat{F}_{Neutral}$ may not exist exactly on F_{face} . In this paper, we use an iterative method to find the neutral face. The method is presented in Figure 4.

The results of expression processing of ten different people are presented in the second line of Figure 5. Error maps are showed in the third line. From the maps, we could find that our method can maintain the rigid parts of the faces. Note that not only can our method remove facial expression, but also it can maintain neutral faces. So in the recognition section, we do not have to

recognize whether the probe face is expressional. Some neutral faces are presented in Figure 6.

Finally, the expression-removed depth images are constructed using $F_{Neutral}$. The size of the depth image is 128×128 .

4 3D face recognition using dual-tree complex wavelet feature

After removing the facial expression, the 3D faces become very similar. Extracting discriminating feature from each face is very important. In this paper, we utilize the dual-tree complex wavelet transform [19,20] to extract expression-removed faces' feature (the size of the face image is 128×128), and six subregions' feature (the six regions are extracted from the face image which size is 384×384 and the size of each region is 128×128). Six feature points are shown in Figure 7A. We used an easy way to find these points. Firstly, we manually defined the six points of a standard face. Then, for each gallery and probe faces, six subregions of size 9×9 which centroids are the same as the standard face are found. Finally, the shape index value [21] refines the six feature points. The local maxima refine landmarks for point 2 and point 5, while the local minima refine landmarks for points 1, 3, 4, and 6. Thus, the six subregions of size 128×128 which centroids are the six refined feature points are defined.

In the training section, we use all the 943 faces in FRGC 1.0 for training. First of all, we extract the four-level magnitude subimages of each training face. Subsequently, we vectorize the six magnitude subimages into a large vector (the dimension is 384), and then we utilize LDA [2] to learn the discriminant subspace and record the

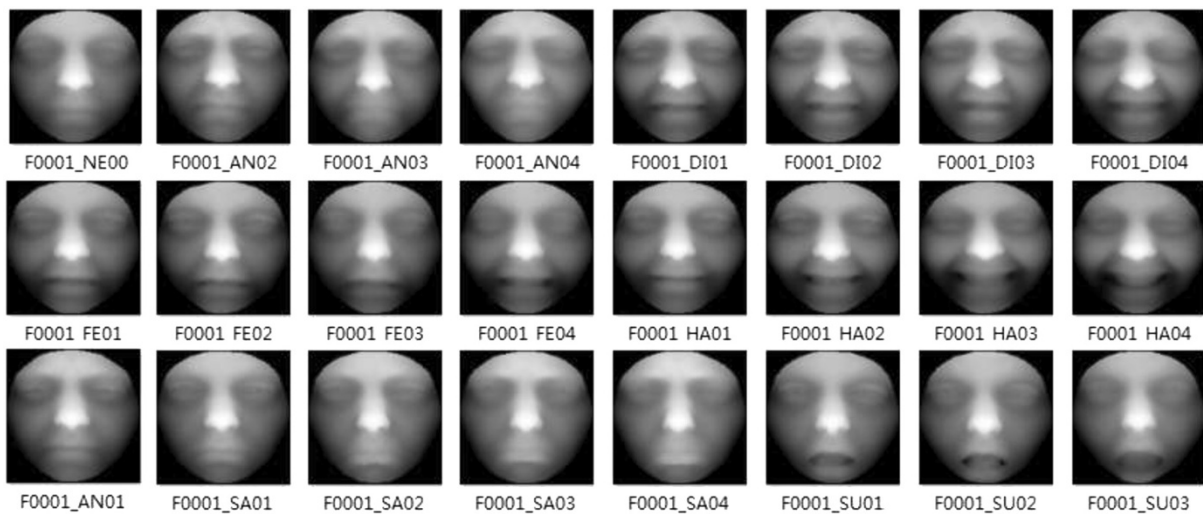
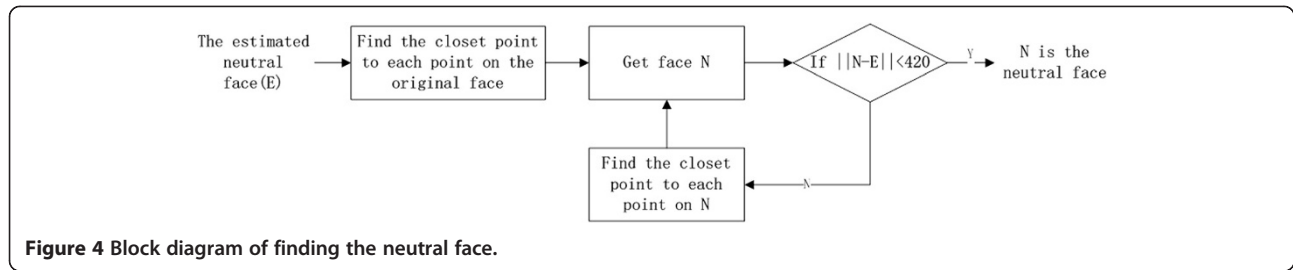


Figure 3 The reference face and expression faces. The first face is the reference face, while the others are the 23 expressional faces which are used to establish expressional space.



transformation matrix. Secondly, we extract the six subregions' four-level magnitude subimages using DT-CWT and vectorize them into a large vector (the dimension is 2,304) and utilize LDA to learn the subspace too. Finally, we get all the gallery faces' two features using DT-CWT and their transformation matrix, respectively.

In the testing section, we get all the probe faces' two features by using DT-CWT and the two transformation matrices, respectively. Cosine distance is used to establish similarity matrix S_1 and S_2 . After this, we normalize them using function (9).

$$S'_{rc} = \frac{S_{rc} - \min(S_r)}{\max(S_r) - \min(S_r)} \quad (6)$$

In the function, S_{rc} represent an element of similarity matrix S_1 and S_2 (at row r and column c), S_r is the elements of S_1 and S_2 at row r , and S'_{rc} denotes the similarity normalized S_{rc} . Then, the final similarity matrix is established by a simple sum rule $S = S_1 + S_2$. Recognition is achieved by the nearest neighbor classifier.

5 Results and analysis

We perform our experiments on the Bosphorus database [22] and the FRGC [23] 3D face database.

The Bosphorus database consists of 105 subjects in various poses, expressions, and occlusion conditions. Eighteen subjects have beard/moustache and short facial hair is available for 15 subjects. The majority of the subjects are aged between 25 and 35 years. There are 60

men and 45 women in total, and most of the subjects are Caucasian. Also, 27 professional actors/actresses are incorporated in the database. Up to 54 face scans are available per subject, but 34 of these subjects have 31 scans. Thus, the number of total face scans is 4,652.

FRGC v1 contained 943 3D faces, while FRGC v2 contained 4,007 3D faces of 466 persons. The images were acquired with a Minolta Vivid 910. The Minolta 910 scanner uses triangulation with a laser stripe projector to build a 3D model of the face. The 3D faces are available in the form of four matrices, each of size 640×480 . The data consists of frontal views. Some of the subject has facial hair, but none of them is wearing glasses. The 2D faces are corresponding to their respective 3D face. In FRGC v2, 57% are male and 43% are female. The database was collected during 2003 to 2004. In order to evaluate the robustness of our method against expression variations, we classified 1,648 faces with expression as the non-neutral dataset (411 persons), while 2,359 neutral faces as the neutral dataset (422 persons). The number of the neutral dataset and the non-neutral dataset is not equal because some people in FRGC v2 contained only one face. We use 'N' which represents for neutral, 'E' which indicates for non-neutral, and 'A' which stands for all in the following of the paper.

5.1 Experiments on Bosphorus database

Firstly, to evaluate the performance of the nose tip detection method, we test our method on the Bosphorus database. The results of data preprocessing of the first

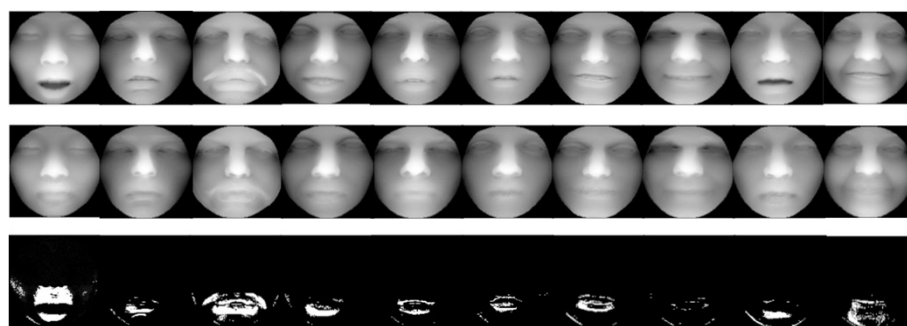


Figure 5 The results of expression processing of ten different people. The first line shows the original depth images. The second line is the expression processing results. The last line is the binary process results of the error maps.

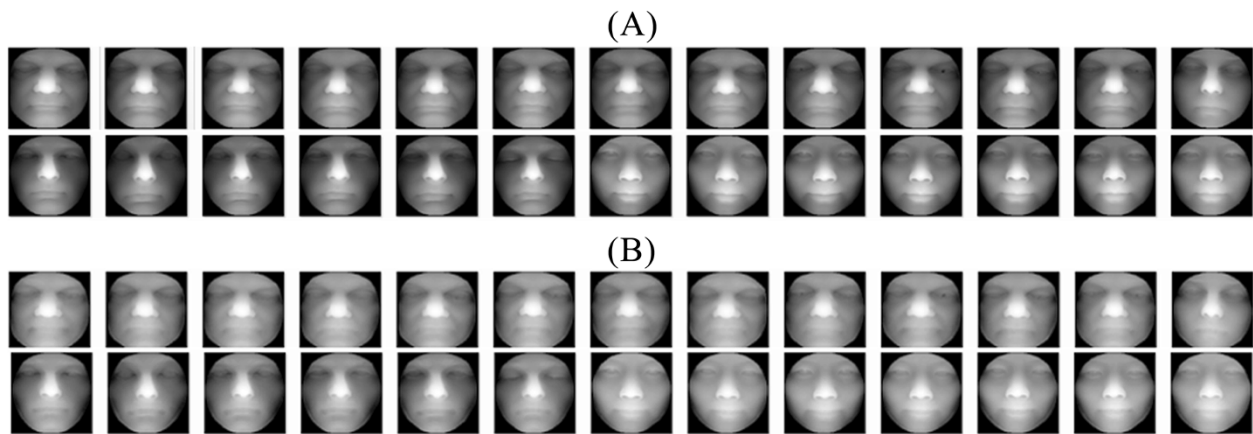


Figure 6 Some neutral faces. Our method can maintain neutral faces.

person are presented in Figure 8. From the figure, we can find that our method can deal with expressional face and posed face which angle is less than 30° , but it cannot find the nose tip of the big angled face ($\pm 45^\circ$ and $\pm 90^\circ$), because most part of the face is missing.

To further confirm the effectiveness of the proposed expression processing approach, we perform experiments on nearly frontal faces (those poses are less than 30°) with expression changes and occlusions. We compare the original faces and the expression-removed faces using leave-one-out method. We extract the DT-CWT feature and then use LDA to finish recognition. FRGC v1 is used for training LDA subspace. Receiver operating characteristic (ROC) curves of the experiment were presented in Figure 9. From the figure, we can find that the facial expression-removed faces performed better than the original faces.

5.2 Experiments on FRGC

5.2.1 Comparison with original mouths

Dealing with open mouth has been a serious topic in 3D face recognition, and a number of researchers have been working on it. We expect that our method in correctly establishing correspondence between an open mouth and estimated neutral component can greatly improve 3D face recognition.

As a first set of experiments, we test our algorithm on the mouth area of FRGC v2. As the experimental protocol, we constructed the gallery set containing the first neutral face for each subject and the remaining ones made up of the probe set. We compare the expression-removed mouths with the original mouths using the PCA method. The recognition rate of using the original mouths is 52.95%, while the recognition rate of using the expression-removed mouths is 69.5%. We could find that the expression-removed mouths contain more identity information than the original mouths.

5.2.2 Comparison with original faces

Then, for the purpose of evaluating the performance of the expression processing method, we compare the expression-removed faces with the original faces using the Gabor feature [24] and DT-CWT feature of the whole depth image. We finished four experiments which contained the neutral vs. neutral experiment, neutral vs. non-neutral experiment, all vs. all experiment, and ROCIII experiment. In the all vs. all experiment, every image of FRGC v2 is matched with all remaining others. It resulted 16,052,042 combinations. Similarly, in the neutral vs. neutral experiment, every image of the neutral database is matched with all remaining others and it resulted 5,562,522 combinations. In the neutral vs. non-neutral experiment, the gallery images come from the neutral dataset and the probe entries come from the expression dataset. In the ROCIII experiment, the gallery

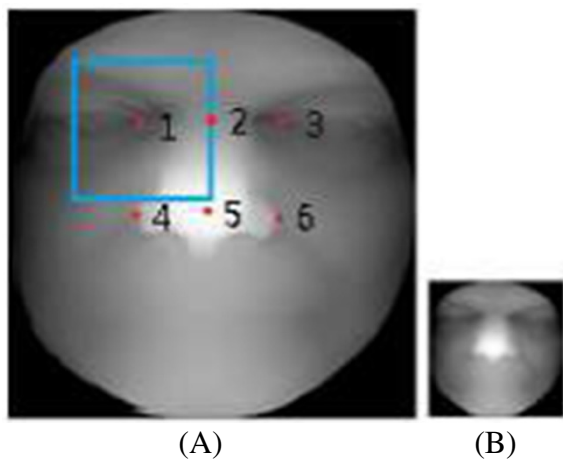


Figure 7 Two depth images with different sizes which are used in the method. (A) Size of 384×384 . (B) Size of 128×128 .

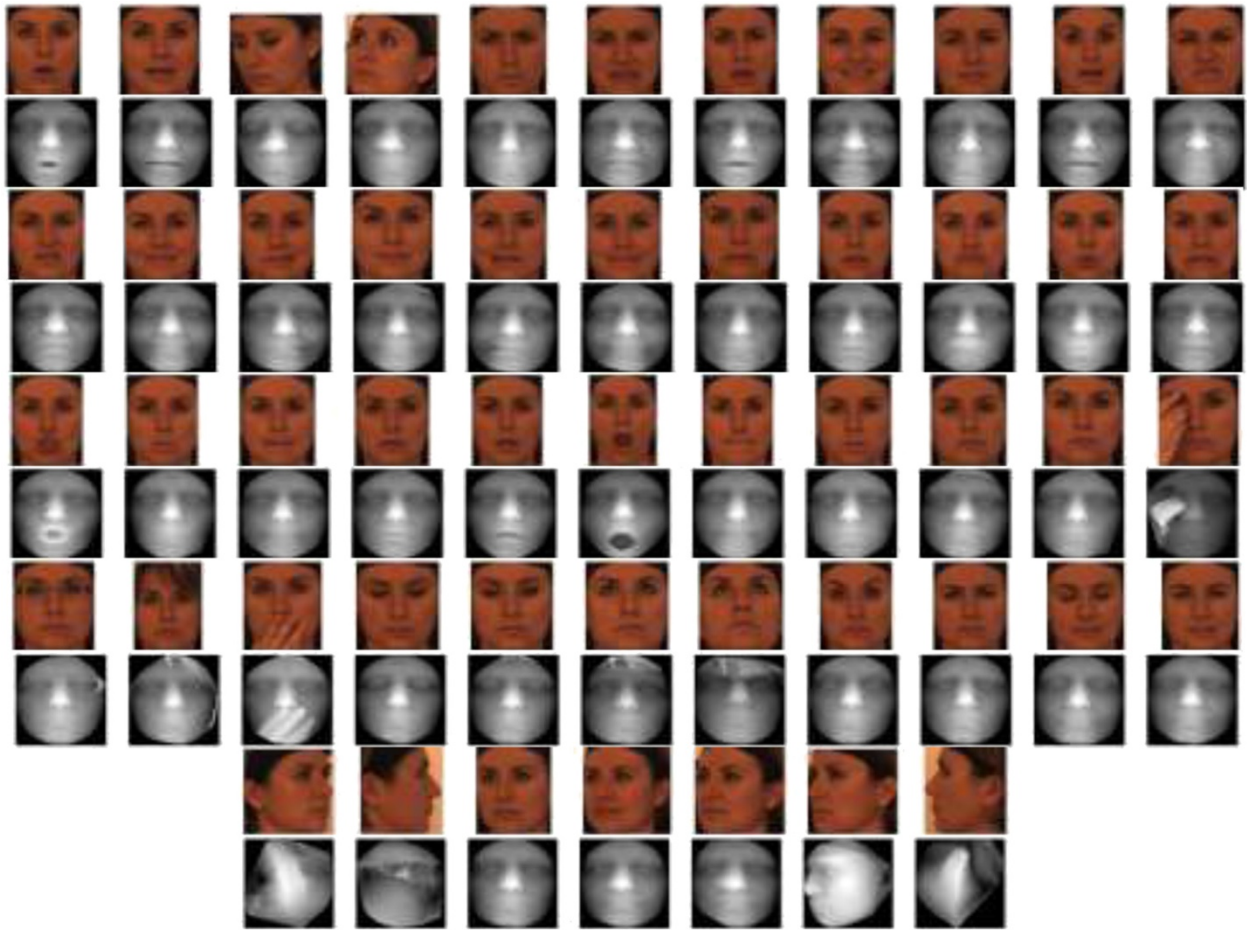


Figure 8 The results of data preprocessing of the first person in the Bosphorus database.

images come from the Fall 2003 semester, while the probe entries come from the Spring 2004 semester.

From Table 1, we can see that the facial expression-removed faces performed better than the original faces. From the N vs. E experiment, we can find that the

expression-removed faces are more useful for face recognition. Using the facial expression-removed faces achieved 11.7% higher recognition rate and 12.29% higher verification rate at 0.001 FAR than using the original faces. Meanwhile, we can find that the DT-CWT

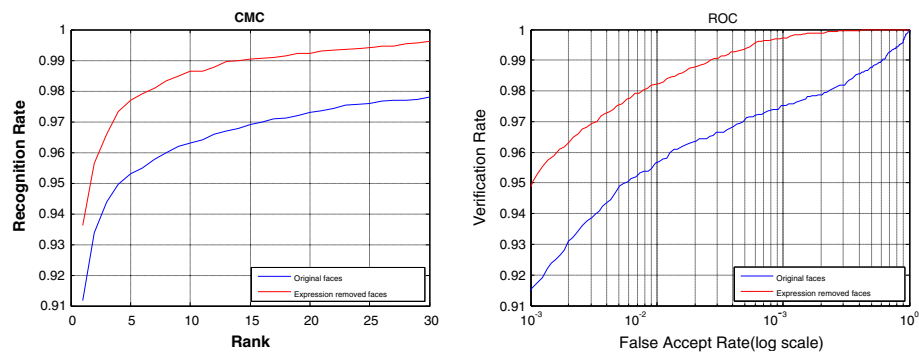


Figure 9 CMC and ROC curves. CMC curves and ROC curves using the DT-CWT feature of the original faces and the DT-CWT feature of the expression-removed faces.

Table 1 Facial expression-removed faces compared with original faces using Gabor feature and DT-CWT feature of whole depth image

Experiments	Feature	Rank one recognition rate		Verification rate at 0.001 FAR	
		Original faces	Facial expression-removed faces	Original faces	Facial expression-removed faces
N vs. N	Gabor	94.57%	95%	96.72%	97.31%
	DT-CWT	98.39%	98.46%	99.09%	99.22%
N vs. E	Gabor	59.92%	67.23%	63.03%	69.88%
	DT-CWT	78.8%	90.5%	81.25%	93.54%
A vs. A	Gabor	83.73%	86.56%	86.79%	90.15%
	DT-CWT	91.9%	95.78%	94.28%	98.03%
ROCI	Gabor	-	-	81.27%	86.61%
	DT-CWT	-	-	89.68%	95.09%

feature is more effective than the Gabor feature for 3D face recognition.

5.2.3 ROC and CMC of our method

In this section, we employed two different scenarios for the experiments: identification and verification. Four experiments were performed as those in Section 5.1.

CMC curves of the four experiments were presented in Figure 10, while ROC curves of the four experiments were presented in Figure 11. The performance of the feature extracted from the whole face using DT-CWT and LDA, the feature extracted from six subregions using DT-CWT and LDA, and the fusion of the two are shown in each figure. As is seen, combining the DT-CWT

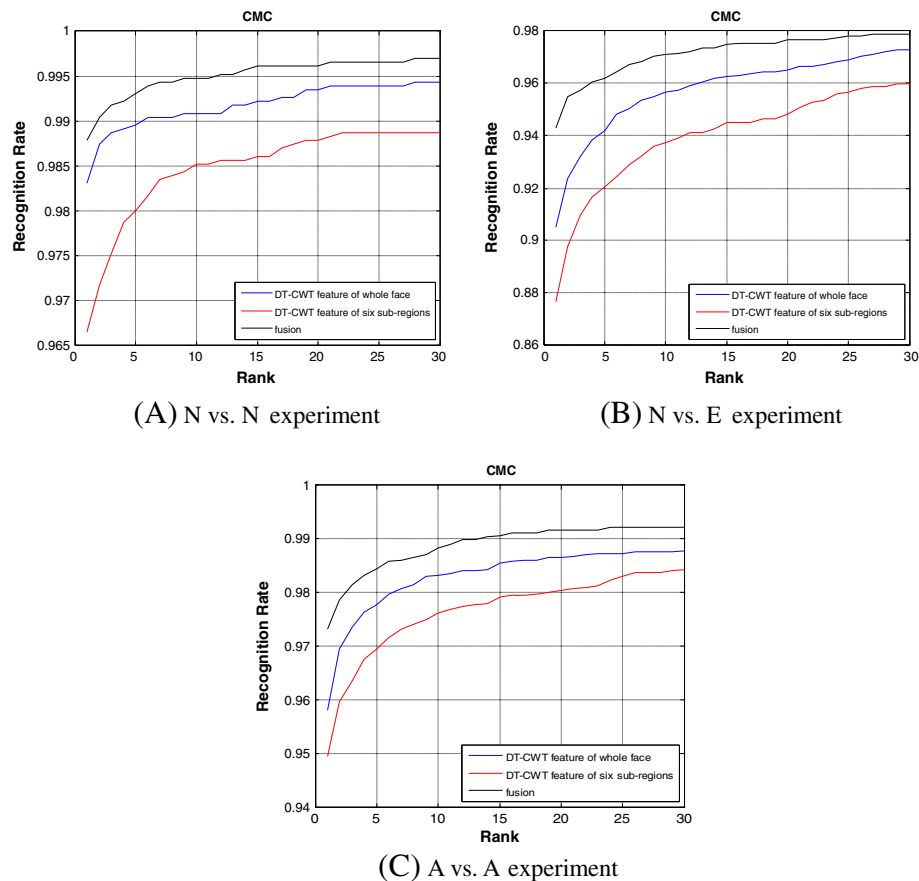


Figure 10 CMC curves. CMC curves using the feature of the whole face, the feature of the six subregions, and the fusion of the two.

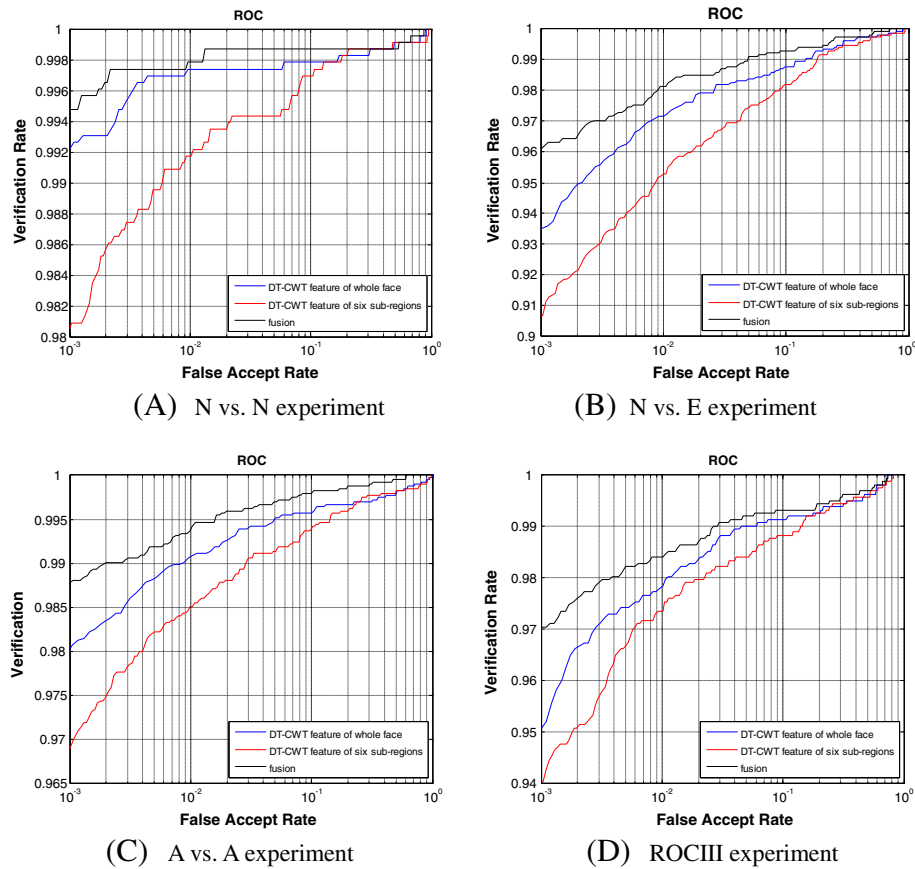


Figure 11 ROC curves. ROC curves using the feature of the whole face, the feature of the six subregions, and the fusion of the two.

feature and the six subregions' feature can improve the recognition performance even further.

5.2.4 Comparisons with other methods

Here, we compared our method with that of the state-of-the-art methods using the fusion results. Table 2 shows the verification results for state-of-the-art methods on the FRGC database as reported in the literature.

Also, the verification rate of our method is shown in Table 2. The performances of A vs. A and ROCIII experiments were slightly lower but still closed to the best.

6 Conclusions

We presented an automatic method for 3D face recognition. We used an improved detection method to correct the pose of the face. We showed that the proposed method could correct posed face which angle is less than 30° .

We also proposed a 3D facial expression processing method, which was based on sparse representation. It could abstract the neutral component from a dictionary which is the combination of neutral and expressional

spaces and enhance the recognition rate. Our method could deal opened mouth and expression of grin. We showed that the estimated neutral faces which are extracted from the expression faces are familiar with that extracted from their corresponding neutral face.

Table 2 Verification rate comparison with the state-of-the-art methods at 0.001 FAR

	N vs. N	N vs. E	A vs. A	ROCIII
Mian [25]	99.4%	-	86.6%	-
Maurer [26]	97.8%	-	87%	-
Cook [27]	-	-	92.31%	92.01%
Faltemier [4]	-	-	93.2%	94.8%
Alyuz [9]	-	-	-	85.64%
Zhang [8]	98.3%	89.5%	-	-
Wang [5]	-	-	98.13%	98.04%
Berretti [7]	97.7%	-	81.2%	-
Kakadiaris [10]	-	-	-	97%
Drira [11]	-	-	93.96%	97%
Mohammadzade [12]	-	-	99.6%	99.2%
Our method (training set: FRGC v1)	99.53%	96.25%	98.86%	97.05%

Then, the facial representation which contained the whole facial feature and the six subregions' feature extracted by DT-CWT were gotten. Holistic and local feature could represent a 3D face more effective for the recognition. Finally, LDA was used to enhance the accuracy of the recognition.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported partly by the National Natural Science Foundation of China (61172128), the National Key Basic Research Program of China (2012CB316304), the New Century Excellent Talents in University (NCET-12-0768), the Fundamental Research Funds for the Central Universities (2013JBM020, 2013JBZ003), the Program for Innovative Research Team in the University of Ministry of Education of China (IRT201206), the Beijing Higher Education Young Elite Teacher Project (YETP0544), the National Natural Science Foundation of China (61403024), and the Research Fund for the Doctoral Program of Higher Education of China (20120009110008, 20120009120009).

Received: 1 April 2014 Accepted: 6 November 2014
Published: 25 November 2014

References

1. C Zhong, Z Sun, T Tan, Robust 3D face recognition using learned visual codebook, in *Proceedings of IEEE Conference on Pattern Recognition* (Minneapolis, 2007), pp. 17–22
2. C Zhong, Z Sun, T Tan, Learning efficient codes for 3D face recognition, in *Proceedings of 15th IEEE International Conference on Image Processing* (San Diego, 2008), pp. 1928–1931. 12–15 Oct
3. KI Chang, KW Bowyer, PJ Flynn, An evaluation of multi-modal 2D +3D face biometrics. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(4), 619–624 (2005)
4. TC Faltmier, KW Bowyer, PL Flynn, A region ensemble for 3-D face recognition. *IEEE Trans. Inf. Forensics Secur.* **3**(1), 62–73 (2008)
5. Y Wang, J Liu, X Tang, Robust 3D face recognition by local shape difference boosting. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(10), 1858–1870 (2010)
6. H Lee, A Battle, R Raina, AY Ng, Efficient sparse coding algorithms. *Adv. Neural. Inf. Proc. Syst.* **19**, 801 (2007)
7. S Berretti, AD Bimbo, P Pala, 3D face recognition using iso-geodesic stripes. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(12), 2162–2177 (2010)
8. G Zhang, Y Wang, Robust 3D face recognition based in resolution invariant features. *Pattern Recognit. Lett.* **32**(7), 1009–1019 (2011)
9. N Alyuz, B Gökberk, L Akarun, Regional registration for expression resistant 3-D face recognition. *IEEE Trans. Inf. Forensics Secur.* **5**(3), 425–440 (2010)
10. A Kakadiaris, G Passalis, G Toderici, MN Murtuza, Y Lu, N Karampatziakis, T Theoharis, Three-dimensional face recognition in the presence of facial expressions: an annotated deformable model approach. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4), 640–649 (2007)
11. H Drira, BB Amor, A Srivastava, M Daoudi, R Slama, 3D face recognition under expressions. *Occlusions Pose Variat.* **35**(9), 2270–2283 (2013)
12. H Mohammadzade, D Hatzinakos, Iterative closest normal point for 3D face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(2), 381–397 (2013)
13. B Amberg, R Knothe, T Vetter, Expression invariant 3D face recognition with a Morphable Model, in *International Conference on Automatic Face & Gesture Recognition* (Amsterdam, 2008), pp. 1–6. 17–19 Sept
14. PN Belhumeur, JP Hespanha, DJ Kriegman, Eigenfaces vs. fisherface: recognition using class special linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 711–720 (1997)
15. X Li, T Jia, H Zhang, Expression-insensitive 3D face recognition using sparse representation, in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Miami, 2009), pp. 2575–2582. 20–25 June
16. X Wang, Q Ruan, Y Jin, G An, Expression robust three-dimensional face recognition based on Gaussian filter and dual-tree complex wavelet transform. *J. Intell. Fuzzy Syst.* **26**, 193–201 (2014)
17. PJ Besl, ND McKay, A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992)
18. D Donoho, For most large underdetermined systems of linear equations the minimal ℓ_1 -norm solution is also the sparsest solution. *Commun. Pure Appl. Math.* **59**(6), 797–829 (2006)
19. IW Selesnick, RG Baraniuk, NG Kingsbury, The dual-tree complex wavelet transform. *IEEE Signal Proc. Mag.* **22**(6), 123–151 (2005)
20. C Liu, D Dai, Face recognition using dual-tree complex wavelet features. *IEEE Trans. Image Process.* **18**(11), 2593–2599 (2009)
21. JJ Koenderink, AJ van Doorn, Surface shape and curvature scales. *Image Vision Comput.* **10**(8), 557–565 (1992)
22. A Savran, N Alyüz, H Dibeklioglu, O Çeliktutan, B Gökberk, B Sankur, L Akarun, Bosphorus Database for 3D face analysis, in *Workshop on Biometrics and Identity Management*, 2008, pp. 47–56
23. PJ Phillips, P Flynn, T Scruggs, KW Bowyer, J Chang, K Hoffman, J Marques, J Min, W Worek, Overview of the Face Recognition Grand Challenge, in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1 (San Diego, 2005), pp. 947–954. 20–25 June
24. JP Jones, LA Palmer, An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* **27**, 1233–1258 (1987)
25. AS Mian, M Bennamoun, R Owens, An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(11), 1927–1943 (2007)
26. T Maurer, D Guignonis, I Maslov, B Pesenti, A Tsaregorodtsev, D West, G Medioni, Performance of Geomatrix ActiVeID TM 3D face recognition engine on the FRGC data, in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (San Diego, 2005), p. 154. 20–25 June
27. J Cook, M Cox, V Chandran, S Sridharan, Robust 3D face recognition from expression categorisation, *ICB 2007. LNCS* **4642**, 271–280 (2007)

doi:10.1186/1687-5281-2014-51

Cite this article as: Wang et al.: Three-dimensional face recognition under expression variation. *EURASIP Journal on Image and Video Processing* 2014 **2014**:51.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com