

ANALYSIS OF RANGE IMAGES USED IN 3D FACIAL EXPRESSION RECOGNITION SYSTEMS

Xiaoli LI, Qiuqi RUAN, Gaoyun AN, Yi JIN

Institute of Information Science

Beijing Jiaotong University, Beijing 100044, China

✉

*Beijing Key Laboratory of Advanced Information Science and Network Technology
Beijing 100044, China*

e-mail: {09112087, qqruan, gyan, yjin}@bjtu.edu.cn

Abstract. With the creation of BU-3DFE database the research on 3D facial expression recognition has been fostered; however, it is limited by the development of 3D algorithms. Range image is the strategy for solving the problems of 3D recognition based on 2D algorithms. Recently, there are some methods to capture range images, but they are always combined with the preprocess, registration, etc. stages, so it is hard to tell which of these generated range images is of higher quality. This paper introduces two kinds of range images and selects different kinds of features based on different levels of expressions to validate the performances of proposed range images; two other kinds of range images based on previously used nose tip detection methods are applied to compare the quality of generated range images; and finally some recently published works on 3D facial expression recognition are listed for comparison. With the experimental results, we can see that the performances of two proposed range images with different kinds of features are all higher than 88% which is remarkable compared with the most recently published methods for 3D facial expression recognition; the analysis of the different kinds of facial expressions shows that the proposed range images do not lose primary discriminative information for recognition; the performances of range images using different kinds of nose tip detection methods are almost the same what means that the nose tip detection is not decisive to the quality of range images; moreover, the proposed range images can be captured without any manual intervention what is eagerly required in safety systems.

Keywords: 3D facial expression recognition, range image, nose tip detection, spatial feature, Fourier transform

Mathematics Subject Classification 2010: 65-D18

1 INTRODUCTION

Facial expression recognition is one of the most important parts in psychology research. Person's mental activities can be sensed by his facial expressions, thus it is significant in security systems. Facial expression recognition will also play an important role in changing the traditional human-machine interactive mode, which is crucially desired in entertainment places, etc. All these potential applications inspired numbers of researchers to facial expression recognition.

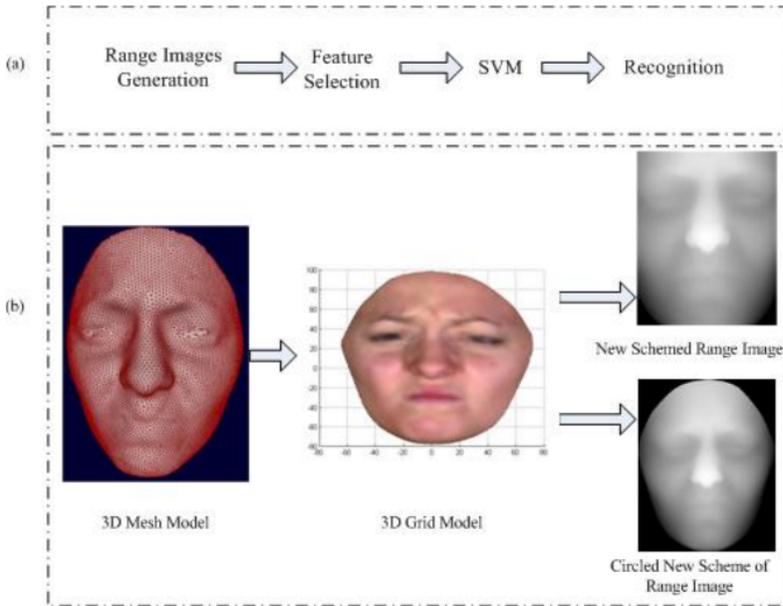


Figure 1. a) The framework of the recognition system; b) Details of range images generation

Facial expression recognition, which is traditionally based on 2D static images [1, 2, 3, 4] or 2D image sequences [5, 6, 7, 8], has been researched for decades. Even though, there are still many problems, such as the change of illumination and large pose variations which are the instinct attributions of 2D images. Thus, some related work [9, 10, 11, 12] has explored these problems into higher dimensional images. With the creation of BU-3DFE database [13] in 2006, the research for 3D facial expression recognition is fostered and has got a great improvement. The following list illustrates the most recent work based on BU-3DFE database:

- Wang et al. [14] performed the very first work of 3D facial expression recognition on BU-3DFE database. They manually located 64 fiducial points to construct

7 expression local regions and extracted 12 primitive facial features for recognition.

- Hamit and Hasan [15, 16] manually labeled 11 feature points to generate 6 distance vectors for recognition, and a previously trained neural network (PNN) was taken as the classifier.
- Tang and Thomas [17] extracted a pool of candidate features based on the points predefined in BU-3DFE database, and from the pool they selected 24 normalized Euclidean distances as the “best” features using the relative entropy. Also, in [18] they performed person and gender independent facial expression recognition based on properties of the line segments connecting certain 3D facial feature points.
- Venkatesh et al. [19] proposed modified PCA for recognition based on shape information which was generated by mapping uniformly sampled contours onto 3D facial models.
- Umut et al. [20] obtained a complete set of distance vectors based on the points provided by features points predefined in BU-3DFE database. And NSGA II was used to determine the optimal set of facial features.

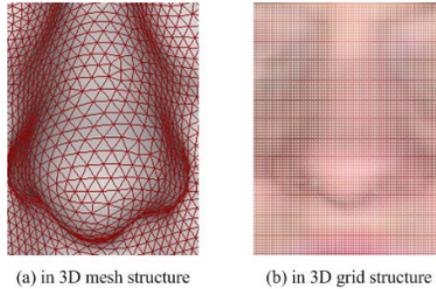


Figure 2. The nose region representation in different structures

All these mentioned works are based on 3D facial models, which can solve intrinsic problems of 2D images. However, as the 3D techniques are newly developed, the algorithms for 3D facial expression recognition are immature. We consider solving this problem with 2D methods. To use 2D methods, the 3D facial models should be projected into 2D space first. Therefore, 2D range images are the first choice to solve this problem. This strategy has been applied in [21] where 2D range images were used to compute SIFT descriptors based on extracted facial landmarks. Meanwhile, 2D range images are recently used for 3D face recognition in [22, 23, 24, 25], where the 3D problem is well resolved. In [22], the range images were used to detect the facial landmarks to solve the problem in face segmentation; the range images were used for registration as the similarity measure based on a simulated annealing-based approach with the surface interpenetration measure in [23]; and in [24], the

range images were used to get the linear combination of the example faces to get the warping matrix which was used for face recognition.

With this strategy, the quality of range images plays a decisive role in 3D facial expression recognition, so researchers pointed to get higher qualified range images. As the location of the nose tip determines the region of face and the quality of range images, some of the researchers tended to improve the nose tip detection method; however, most of their work [23, 25] are combined of several preprocessing, registration, calculation, etc. stages in the recognition system; the final recognizing results are the expression of the combination, so we cannot say whether the nose tip detection is effective to get the qualified range images. This paper is to contribute to analysis of the range images based on different nose tip detection methods, so previous preprocess of the 3D models is avoided as much as possible. Therefore, all the 3D models are assumed to be well-posed or almost frontal posed.

The models captured in BU-3DFE database [13] all provided the orientation with respect to the frontal projection plane which can be seen as frontal pose. Therefore, this paper uses BU-3DFE database to transform the models into a new scheme of range images directly; to verify the performance of a new scheme of range images, different traditional features are selected for comparison and different levels of expressions are considered; to compare the performances of different methods for detecting nose tip, two previous methods for nose tip detection are tested in the experimental section.

The framework of 3D facial expression recognition system in this paper is shown in Figure 1 a). It can be mainly divided into three steps: firstly, the range images are generated based on original 3D mesh facial models; secondly, distinct facial expression features are extracted based on proposed range images; finally SVM (Support Vector Machine) is taken as the classifier for recognition. Among these three steps, the first step, range images generation, is the key point that is emphasized in this paper, and it is embodied in Figure 1 b). Also, its introduction will be detailed in Section 2. Section 3 will introduce the selection of spatial features, and Section 4 will give the experimental results. Finally, conclusions are put in Section 5.

2 GENERATION OF A NEW SCHEME OF RANGE IMAGES

To analyse the performances of range images, a new scheme of range images are generated. Recently, many 3D facial databases [13, 26, 27] are created, but only BU-3DFE (Binghamton University 3D Facial Expression) database [13] is specially designed for 3D facial expression, and the face models captured in BU-3DFE database can be seen as frontal pose which satisfies our requirements. Thus, this paper merely considers the models in BU-3DFE database. To generate the new scheme of range images, the mesh models are converted into 3D grid models first.

Compared with 3D mesh facial models, we can get numerous advantages of 3D grid models: grid models can be seen as matrices which can be easily obtained and

applied to some traditional methods; it is much easier to get the corresponding points in a matrix, thus the search algorithms can be avoided; especially, range images can be directly constructed using 3D grid models, and the points between them are one-to-one correspondent.

Each node of 3D mesh facial models in BU-3DFE database [13] contains both coordinate information $\{xyz\}$ and color information $\{RGB\}$, so the node can be denoted as $\{xyzRGB\}$. The 3D mesh facial models are converted into 3D grid models in the following steps:

1. Locate the tip of nose. Not using the complex algorithms taken in [19, 28, 29] to detect the nose tip, we just assume that z value of nose tip is the highest in $-1/2^{\text{th}}$ to $1/2^{\text{th}}$ mask of central portion.
2. Shift all the other nodes to the new coordinate space by subtracting the nose tip.
3. Create a grid $A_{m \times n}$ with a uniform space (We set the space of 0.25 in our experiments [30]). Where m is the distribution range of x values, and n is the distribution range of y values.
4. Using the Qhull algorithm [31], obtain z values, cubically interpolated over the nodal coordinates $\{xyz\}$.
5. Similar to the z values, we can also get the R, G, B values in grid $A_{m \times n}$ based on the Qhull algorithm, cubically interpolated over the nodal coordinates $\{xyR\}, \{xyG\}, \{xyB\}$ respectively.

In this manner, we can obtain both the coordinate information and color information for each grid point. So the 3D grid facial model $A_{m \times n \times 6}$ is achieved, and each entry contains $\{xyzRGB\}$ information. The nose region both in 3D mesh structure and 3D grid structure is shown in Figure 2, and samples of 3D grid facial models corresponding to six basic expressions are shown in Figure 3.

Expression	AN	DI	FE
3D Grid Model			
Expression	HA	SA	SU
3D Grid Model			

Figure 3. 3D grid models corresponding to six basic facial expressions. AN-angry, DI-disgust, FE-fear, HA-happy, SA-sad, SU-surprise.

Both coordinate information and color information in the 3D grid model $A_{m \times n \times 6}$, or just part of the information can be chosen to get the range images. However, as the coordinate information is more discriminative than color information [19], this paper only considers coordinate information $\{xyz\}$ of each entry in $A_{m \times n \times 6}$ that can be noted as $A_{m \times n \times 3}$.

Expression	New Schemed Range Images	Circled New Scheme of Range Images	Highest z Range Images	Central Profiled Range Images
AN				
DI				
FE				
HA				
SA				
SU				

Figure 4. Representation of different kinds of range images corresponding to six basic facial expressions. AN-angry, DI-disgust, FE-fear, HA-happy, SA-sad, SU-surprise.

According to the values in $A_{m \times n \times 3}$, we set the x and y grid values as the image grid and z values as the gray intensity of the image, so the new scheme of range images of six expressions are projected and shown in Figure 4 column 2, naming “new schemed range images”. To compare with the most recently used elliptical range images, the “new schemed range images” are then cropped with the elliptical

nose region (The nose tip is used as the center of the circle and then the circle is uniformed as the same size.) which are shown in Figure 4 column 3, naming “circled new scheme of range images”.

3 FEATURES SELECTED IN SPATIAL DOMAIN

To validate the two proposed range images for facial expression recognition, features both in spatial domain and in frequency domain are considered. As FFT is well-known and the features in frequency domain are just the application of FFT on the range images, here only the two kinds of spatial features are introduced in the following subsections.

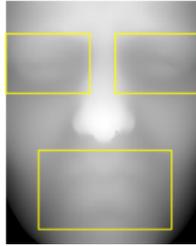


Figure 5. Selected local regions including eyes region and mouth region

3.1 Global Features Selected in Spatial Domain

The gray intensity of the whole range image is considered to get the global spatial features. As the scale of each range image is 401×501 , we extract the global spatial feature as a 401×501 matrix, and reshape it as a 1×200901 vector for recognition.

3.2 Local Features Selected in Spatial Domain

As human faces may be partly disguised with adornments, such as glasses, scarves, mask, etc., the global information of facial expressions are hardly to be obtained. Moreover, with the privacy requirements, the influence of local expression regions is desired to investigate. The information of facial expressions mainly focuses on the local regions of eyes and mouth [32, 33], therefore, we merely consider these primary local regions.

Since the original facial models in BU-3DFE database are well posed, the range images are well uniformed. To be simple, fixed positions in the range images are chosen as the eyes and mouth regions. With the visual impression, we select the eyes region and mouth region as 124×172 and 164×272 respectively, which are shown in Figure 5. With the chosen local regions, the spatial features for each range image can be reduced as a 1×87264 vector for recognition.

4 EXPERIMENTAL RESULTS AND DISCUSSION

The BU-3DFE database contains 100 subjects (56% female, 44% male). Each subject performed neutral and six basic expressions. With the exception of the neutral expression, each of the six prototypic expressions includes four levels of intensity. Referred to the recent work on 3D facial expression recognition based on BU-3DFE database, only higher two level expressions are considered in the following experiments except the discussion of different level expressions. In the following experiments, all the training and testing subjects are chosen randomly to be gender-independent; referred to the previous work and to be person-independent, 60 subjects in each facial expression are randomly chosen, where 54 subjects as a training set and the rest 6 subjects as a testing set; each experiment is repeated for 10 times to lower the randomness.

4.1 Comparisons among SVMs with Different Kernels

Original SVM is suitable for binary classification. However, it can be easily extended to a multi-class problem using one-against-one approach and one-against-all approach [34]. As the “one-against-one” approach is time consuming, the multi-class SVM classifier using “one-against-all” approach is constructed in this paper. The “one-against-all” SVM classifier schemes on an n -class classification; it considers n binary classifications each of which labels one class as $(+1)$ and all other $n - 1$ classes as (-1) .

The “one-against-all” multi-class SVMs with three different kernels, linear function, Gaussian function, sigmoid function are tested to decide which kernel is used in the following experiments and the global spatial features extracted from “new schemed range images” are considered. Each global spatial feature is 401×501 and concatenated to form a 1×200901 vector for recognizing. The recognition rates (RR%) for six basic facial expressions are listed in Table 1. Obviously, from Table 1 the SVM with Gaussian kernel (88.78%) is seen to perform best, so it is taken as the sole classifier in the following experiments.

SVM kernel	RR%
Linear function	81.33
Gaussian function	88.78
Sigmoid function	84.67

Table 1. Recognition rates based on global spatial features with different kernelled SVM

4.2 Comparisons with the Most Recently Used Range Images

To compare the performances of different methods for detecting nose tip, two previous methods for nose tip detection are tested in this section. Therefore, two kinds of range images are generated:

The first one is generated based on the following nose tip detection method: searching the highest z values in the whole face [23, 24] as the nose tip. Samples of the generated range images are named as “highest z range images” and shown in Figure 4 column 4.

The second one is generated based on the following nose tip detection method: Locate the nose tip based on the central profile which was obtained by the intersection between the symmetry plane and the facial model [25], and the range images are cropped with the elliptical nose region. The formed range images are named as “central profiled range images” and samples of them are shown in Figure 4 column 5.

Model No.	New Schemed Range Images	Circled New Scheme of Range Images	Highest z Range Images	Central Profiled Range Images
F0005				
F0006				
F0028				
M0039				

Figure 6. Certain samples of range images in “Sad” expression formed by different methods. F0005, F0006, F0028 and M0039 are the numbers of facial model set in BU-3DFE database.

According to the steps of generating all these range images, we can observe the main differences focus on the methods for nose tip location. The integrity of facial expression regions obtained by these range images is determined by the accuracy of locating nose tip. From all the sets of generated range images, the information of facial expression in some range images formed by the two previous methods is severely lost. This phenomenon is especially obvious in “Sad” expression where we select certain samples shown in Figure 6.

In Figure 6, we can clearly observe that with the inaccurate located nose tip, the facial expression regions of F0005, F0006, F0028 and M0039 extracted by previous

methods are not entire. Thus, the nose tip detection method used in proposed two range images is verified to be effective.

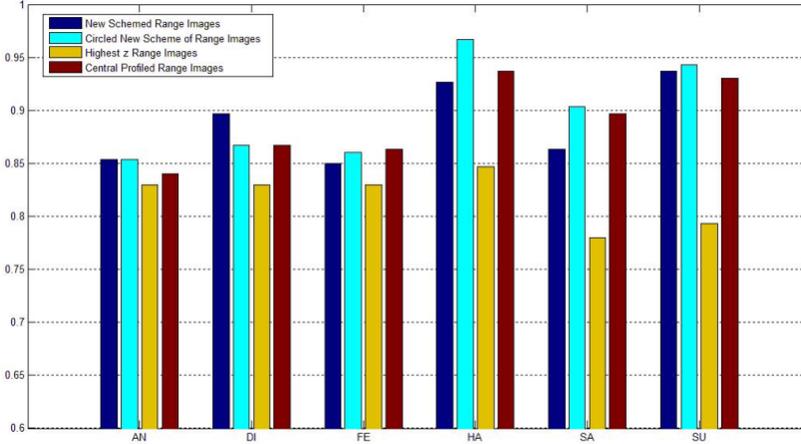


Figure 7. Comparative results of different range images based on extracted global feature. AN-angry, DI-disgust, FE-fear, HA-happy, SA-sad, SU-surprise.

To test the performances of these four kinds of range images, we extract their global spatial features and take “one-against-all” Gaussian kernelled SVM for recognition. The comparative results are shown in Figure 7.

From the comparative results in Figure 7, we can see that in all the six basic expressions, the recognition rates of two proposed range images are much higher than that of “both highest z range images” and “central profiled range images”. So the performance of the global spatial features extracted from proposed range images is much better than that of both “highest z range images” and “central profiled range images”. However, compared with the “central profiled range images”, the performance does not play obviously well. So their average recognition ratios are shown in Table 2. Seen from Table 2, we can clearly get that the result of the “highest z range images” is the lowest; results of other three methods are almost the same, and “circled new scheme of range images” performs best.

Methods	Average Recognition Ratio (RR %)
New schemed range images	88.78
Circled new scheme of range images	89.89
Highest z range images	81.83
Central profiled range images	88.89

Table 2. Average recognition results of range images based on global spatial features

From Table 2, we can see that “circled new scheme of range images” performs a bit better than the “new schemed range images”. However, the “circled new scheme of range images” is cropped with the elliptical nose region which can be changed into other shapes. To be universal, the performances of “new schemed range images” are merely considered to validate the performances in the following experiments.

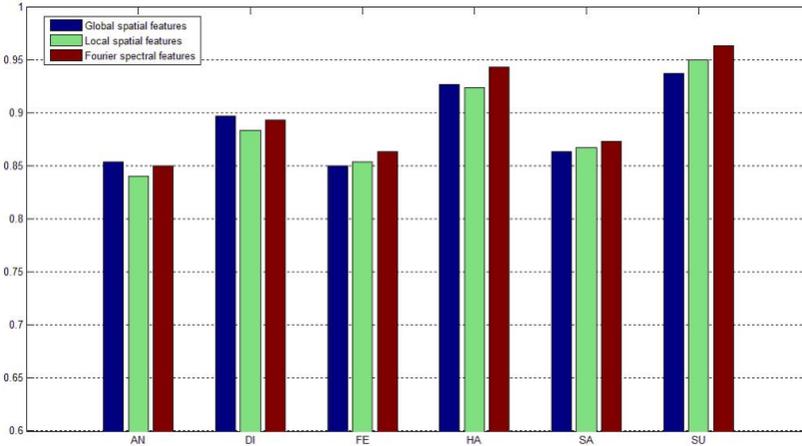


Figure 8. Comparative results of different features extracted from original generated range images. AN-angry, DI-disgust, FE-fear, HA-happy, SA-sad, SU-surprise.

4.3 Comparisons among Features in Spatial Domain and Frequency Domain

To validate the proposed range images for facial expression recognition, we consider discovering their performance of hidden patterns in the Fourier spectral domain [30]. Therefore, both spatial features and Fourier spectral features extracted from original generated range images are considered in this section, and spatial features are considered in both global and local forms. The recognition results for six basic expressions based on these three kinds of features are shown in Figure 8. All the recognition results of each expression in Figure 8 is higher than 85%, especially HA and SU expressions. So, we can conclude that the performances of “new schemed range images” both in spatial domain and frequency domain are considerable.

To be straightforward, the average recognition rates are listed in Table 3. We can clearly get that all the results are higher than 88%, especially the performance of Fourier spectral features, which is almost 90%, these results are remarkable. Meanwhile, we can see that the recognition rates of global spatial features and local spatial features are about the same. These results validate that the discriminative information of facial expressions is concentrated in eyes and mouth regions.

Features	Average Recognition Ratio (RR %)
Global spatial features	88.78
Local spatial features	88.61
Fourier spectral features	89.78

Table 3. Average recognition results of different features

4.4 Analysis Based on Partial Facial Expressions

As the discriminative information of facial expressions is mostly concentrated on eyes and mouth regions [32, 33], and human faces may be partly disguised with adornments, the most valuable local region for recognizing each expression is required for discussion. In addition, benefited from the local region, the storage memory can be reduced sharply, and less time will be used for recognition.

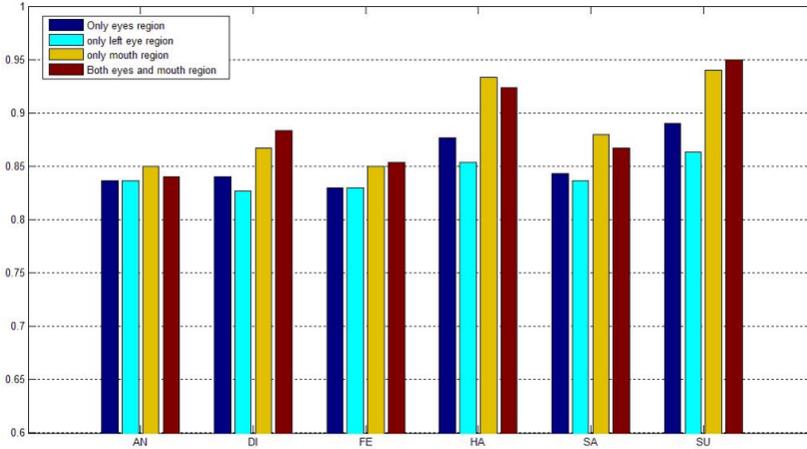


Figure 9. Comparative results of different local features extracted from new schemed range images. AN-angry, DI-disgust, FE-fear, HA-happy, SA-sad, SU-surprise.

To exactly verify performances of new schemed range images, the features in local regions are separately extracted. The results for six basic expressions are shown in Figure 9, and their average recognition rates are listed in Table 4. From the comparable results in Figure 9 and Table 4, we can see that:

1. Data only in the mouth region always performs better than data only in the eyes region and data only in the left eye region. This implies that the information in the mouth region is much more discriminative than that in the eyes region.
2. Data only in the left eye region almost performs the same as data only in the eyes region. The left eye region can provide almost all the expression information in the eyes region, which means that the eyes region is of good symmetry.

3. From Table 4, data in both eyes and mouth region performs better than data only in the eyes region, but worse than data only in the left eye region. This implies that the mouth region plays a more important role than the eyes region does.
4. According to the results of DI, FE and SU expressions in Figure 9, data in both the eyes and mouth region performs better than both data only in the eyes region and data only in the mouth region. The performance of local region features is not only determined by the mouth region, but also the eyes region. The contribution of the eyes region should not be ignored.

According to all the results got from Figure 9 and Table 4, we can conclude that the mouth region plays a decisive role, but the eyes region also affects the final results; the eyes region is of good symmetry. All these conclusions express that new schemed range images possess the attributions of facial expression images, and the projection from 3D space to 2D does not lose their primary discriminative information.

Features	Average Recognition Ratio (RR %)
Only eyes region	85.28
Only left eye region	84.11
Only mouth region	88.67
Both eyes and mouth region (Local spatial features)	88.61

Table 4. Average recognition results of different local features

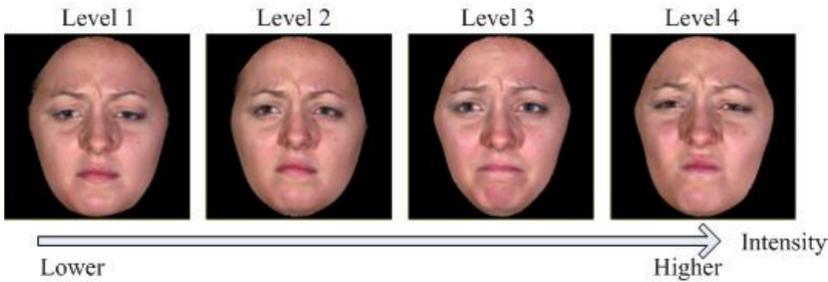


Figure 10. Four level intensive expressions in BU-3DFE database. (Take the subject no. F0001 with “Angry” expression as an example.)

4.5 Comparisons Among Different Intensive Level Expressions

The experiments above only consider the higher intensive two level expressions in BU-3DFE database. At the same time, the lower intensive two level expressions are

also considered to verify the performance of new schemed range images. (These four levels intensive expressions are shown in Figure 10.)

Based on the conditions set in previous experiments, experimental data used in this experiment are with the same design: training and testing sets are selected to be gender-independent and person-independent; 60 subjects in each facial expression are randomly chosen, where 54 subjects as a training set and the rest 6 subjects as a testing set; each experiment is repeated for 10 times to lower the randomness.

Features	Global Spatial Features (RR %)	Local Spatial Features (RR %)
Higher two level expressions	88.78	88.61
Lower two level expressions	86.44	85.61

Table 5. Comparative results of different levels of expressions

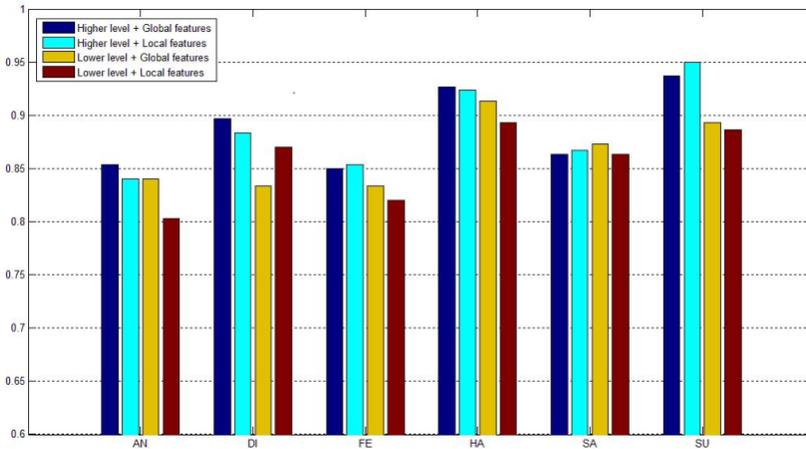


Figure 11. Comparative results of different features based on different levels of expressions. AN-angry, DI-disgust, FE-fear, HA-happy, SA-sad, SU-surprise.

To compare the results clearly, both global and local features are extracted from different levels of new schemed range images. The results for six basic expressions are shown in Figure 11, and their average recognition rates are listed in Table 5. From the comparable results in Figure 11 and Table 5, we can see that:

1. The results of both the global spatial feature and the local spatial features selected from lower two level expressions are higher than 85% and that is remarkable. The range images captured using the original method are reliable.
2. The features selected from higher two level expressions always perform better than that from the lower two level expressions in almost all basic expressions.

This implies that the performance of range images is determined by the original intensity of 3D mesh model expressions.

3. Only in “Sad” expression, the performances of features extracted in the higher and lower two level expressions are almost the same. The variation of “Sad” expression in different levels is smaller relative to other five basic expressions.

4.6 Comparisons with State-of-the-Art Methods

All these recognition results obtained from the two proposed range images are higher than 85 % what is comparable to the recent work in 3D facial expression recognition. The comparative results with the most recently published methods based on BU-3DFE database for 3D facial expression recognition are listed in Table 6. As the previous experiments are based on the higher two level expressions in BU-3DFE database, our results using two higher level expressions are just listed.

We can clearly see that all our results are higher than in other previous works. Moreover, features extracted in these previous works are all based on manually located key points in 3D mesh models; however, our proposed range images are all automatically transformed from the 3D models. Our strategy shown in Figure 1 can be bypassed altogether without any manual intervention, what is crucially required in safety areas. With all these advantages the proposed range images certainly have superiority for 3D facial expression recognition.

Recent Work	Features	RR %
Berretti [21]	SIFT Descriptors	77.54
Venkatesh [19]	Contour & Shape	81.67
Wang [14]	Primitive Surface Label	83.6
Tang [18]	Distance & Slope Information	87.1
Hamit [15]	Distance Vectors	87.8
Umut [20]	Distance Vectors	88.18
original generated range images	Global Spatial Features	88.78
original generated range images	Local Spatial Features	88.61
original generated range images	Fourier Spectral Features	89.78
circle cropped range images	Global Spatial Features	89.89

Table 6. Comparative results with the most recently published work

5 CONCLUSION

The contribution of this paper is an analysis of the range images used in 3D facial expression recognition system. Firstly, two kinds of range images are generated; then different kinds of features based on different levels of expressions are extracted to validate the performances; and two other kinds of range images using previous nose tip detection methods are generated to compare the nose tip detection methods; finally, some recently published works are listed to compare the recognition

results. From the experimental results, we can see that the performances of different features extracted from the proposed range images are all higher than 88% which is remarkable compared to recent work on 3D facial expression recognition and it means the proposed range images do not lose primary discriminative information for recognition; with the well-posed mesh models, the nose tip detection is not so decisive for the quality of range images; moreover, the proposed framework is relatively simple and automatically implemented which is promising for use in practical 3D facial expression recognition systems. All these results can be taken as a directive for 3D facial expression recognition based on range images.

Acknowledgements

This work was supported partly by the National Natural Science Foundation of China (61471032, 61472030, 61403024), National Key Basic Research Program of China (2012CB316304), New Century Excellent Talents in University (NCET-12-0768), the fundamental research funds for the central universities (2013JBZ003), Program for Innovative Research Team in University of the Ministry of Education of China (IRT201206).

REFERENCES

- [1] SHAN, C.—GONG, S.—MCOWAN, P. W.: Facial Expression Recognition Based on Local Binary Patterns: A Comprehensive Study. *Image and Vision Computing*, Vol. 27, 2009, No. 6, pp. 803–816.
- [2] ZHI, R.—RUAN, Q.: Facial Expression Recognition Using Discriminant Locality Preserving Projections. *Journal of Signal Processing*, Vol. 25, 2009, No. 7, pp. 233–237.
- [3] LI, X.—RUAN, Q.—RUAN, C.: Facial Expression Recognition with Local Gabor Filters. 2010 IEEE 10th International Conference on Signal Processing, 2010, pp. 1013–1016.
- [4] LIN, D.: Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network. *Journal of Information Science and Engineering*, Vol. 22, 2006, No. 5, pp. 1033–1046.
- [5] ZENG, Z.—FU, Y.—ROISMAN, G. I. et al.: Spontaneous Emotional Facial Expression Detection. *Journal of Multimedia*, Vol. 1, 2006, No. 5, pp. 1–8.
- [6] ZHU, Y.—DE SILVA, L. C.—KO, C. C.: Using Moment Invariants and HMM in Facial Expression Recognition. *Pattern Recognition Letters*, Vol. 23, 2001, pp. 83–91.
- [7] KUMANO, S.—OTSUKA, K.—YAMATO, J.—MAEDA, E.—SATO, Y.: Pose-Invariant Facial Expression Recognition Using Variable-Intensity Templates. 2007 Asian Conference on Computer Vision, Part I, Vol. 4843, 2007, pp. 324–334.
- [8] BARTLETT, M. S.—LITTLEWORT, G.—FASEL, I.—MOVELLAN, J. R.: Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction. *Computer Vision and Pattern Recognition Workshop*, 2003, pp. 53–58.

- [9] BARTLETT, M.—LITTLEWORT, G.—FRANK, M.—LAINSCSEK, C.—FASEL, I.—MOVELLAN, J.: Recognizing Facial Expressions: Machine Learning and Application to Spontaneous Behavior. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05), San Diego, CA, 2005, Vol. 2, pp. 568–573.
- [10] ZALEWSKI, L.—GONG, S.: Synthesis and Recognition of Facial Expressions in Virtual 3D Views. Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004, pp. 493–498.
- [11] WEN, Z.—HUANG, T.: Capturing Subtle Facial Motions in 3D Face Tracking. Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV), 2003, Vol. 2, pp. 1343–1350.
- [12] YURTKAN, K.—SOYEL, H.—DEMIREL, H. et al.: Face Modeling and Adaptive Texture Mapping for Model Based Video Coding. Conference of Computer Analysis of Images and Patterns, 2005, pp. 498–505.
- [13] YIN, L.—WEI, X.—SUN, Y.—WANG, J.—ROSATO, M. J.: A 3D Facial Expression Database for Facial Behavior Research. 7th International Conference on Automatic Face and Gesture Recognition, 2006, pp. 211–216.
- [14] WANG, J.—YIN, L.—WEI, X.—SUN, Y.: 3D Facial Expression Recognition Based on Primitive Surface Feature Distribution. 2006 IEEE Computer Society Conference on Computer Vision Pattern Recognition, 2006, pp. 1399–1406.
- [15] SOYEL, H.—DEMIREL, H.: 3D Facial Expression Recognition with Geometrically Localized Facial Features. 23rd International Symposium on Computer and Information Science, 2008.
- [16] SOYEL, H.—DEMIREL, H.: Facial Expression Recognition Using 3D Facial Feature Distances. 4th International Conference on Image Analysis and Recognition, 2007, pp. 22–24.
- [17] TANG, H.—HUANG, T. S.: 3D Facial Expression Recognition Based on Automatically Selected Features. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, pp. 1–8.
- [18] TANG, H.—HUANG, T. S.: 3D Facial Expression Recognition Based on Properties of Line Segments Connecting Facial Feature Points. IEEE International Conference on Automatic Face and Gesture Recognition, 2008, pp. 1–6.
- [19] VENKATESH, Y. V.—KASSIM, A. A.—RAMANA MURTHY, O. V.: A Novel Approach to Classification of Facial Expressions from 3D-Mesh Datasets Using Modified PCA. Pattern Recognition Letters, Vol. 30, 2009, No. 12, pp. 1128–1137.
- [20] TEKGUC, U.—SOYEL, H.—DEMIREL, H.: Feature Selection for Person-Independent 3D Facial Expression Recognition Using NSGA-II. 24th International Symposium on Computer and Information Sciences, 2009, pp. 35–38.
- [21] BERRETTI, S.—DEL BIMBO, A.—PALA, P.: A Set of Selected SIFT Features for 3D Facial Expression Recognition. 20th International Conference on Pattern Recognition, 2010, pp. 4125–4128.
- [22] PAMPLONA SEGUNDO, M.—SILVA, L.—BELLON, O. R. P.—QUEIROLO, C. C.: Automatic Face Segmentation and Facial Landmark Detection in Range Images. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 40, 2010, No. 5, pp. 1319–1330.

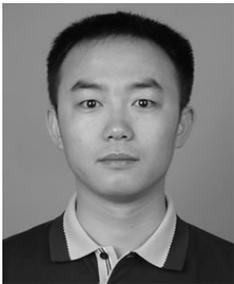
- [23] QUEIROLO, C. C.—SILVA, L.—BELLON, O. R. P.—PAMPLONA SEGUNDO, M.: 3D Face Recognition Using Simulated Annealing and the Surface Interpenetration Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, 2010, No. 2, pp. 206–219.
- [24] ZOU, L.—CHENG, S.—XIONG, Z.—LU, M.—CASTLEMAN, K. R.: 3-D Face Recognition Based on Warped Example Faces. *IEEE Transactions on Information Forensics and Security*, Vol. 2, 2007, No. 3, pp. 513–528.
- [25] WANG, Y.—LIU, J.—TANG, X.: Robust 3D Face Recognition by Local Shape Difference Boosting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, 2010, No. 10, pp. 1858–1870.
- [26] MAURER, T.—GUIGONIS, D.—MASLOV, I.—PESENTI, B.—TSAREGORODTSEV, A.—WEST, D.—MEDIONI, G.: Performance of Geometrix ActiveID TM 3D Face Recognition Engine on the FRGC Data. *IEEE Workshop on Computer Vision and Pattern Recognition*, 2005, pp. 154–160.
- [27] YIN, B.—SUN, Y.—WANG, C.—GE, Y.: BJUT-3D Large Scale 3D Face Database and Information Processing. *Journal of Computer Research and Development*, Vol. 46, 2009, No. 6, pp. 1009–1018.
- [28] MIAN, A. S.—BENAMOUN, M.—OWENS, R.: An Efficient Multi-Modal 2D-3D Hybrid Approach to Automatic Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, 2007, No. 11, pp. 1927–1943.
- [29] SEGUNDO, M. P.—QUEIROLO, C.—BELLON, O. R. P.—SILVA, L.: Automatic 3D Facial Segmentation and Landmark Detection. *14th International Conference on Image Analysis and Processing*, 2007, pp. 431–436.
- [30] VENKATESH, Y. V.—KASSIM, A. K.—RAMANA MURTHY, O. V.: Resampling Approach to Facial Expression Recognition Using 3D Meshes. *20th International Conference on Pattern Recognition*, 2010, pp. 3772–3775.
- [31] BARBER, C. B.—DOBKIN, D. P.—HUHDANPAA, H. T.: The Quickhull Algorithm for Convex Hulls. *ACM Transactions on Mathematical Software*, Vol. 22, 1996, No. 4, pp. 469–483.
- [32] PANTIC, C.—COTTRELL, G.: Representing Face Images for Emotion Classification. *Advances in Neural Information Processing Systems*, 1997, pp. 894–900.
- [33] PARDAS, M.—BONAFONTE, A.: Facial Animation Parameters Extraction and Expression Recognition Using Hidden Markov Models. *Signal Processing: Image Communication*, Vol. 17, 2002, No. 9, pp. 675–688.
- [34] KAZUHIRO HOTTA: Robust Face Recognition Under Partial Occlusion Based on Support Vector Machine with Local Gaussian Summation Kernel. *Image and Vision Computing*, Vol. 26, 2008, No. 11, pp. 1490–1498.



Xiaoli LI is currently a Ph.D. student in Beijing Jiaotong University. Her major is signal processing. Her research interest includes pattern recognition, computer vision, etc.



Qiuqi RUAN received his B.Sc. and M.Sc. degrees from Northern Jiaotong University, Beijing, China, in 1969 and 1981 respectively. From January 1987 to May 1990, he was a Visiting Scholar with the University of Pittsburgh, PA, USA, and the University of Cincinnati, OH, USA. He is currently Professor and Doctorate Supervisor with the Institute of Information Science, Beijing Jiaotong University, Beijing, China. He has published four books and over 100 papers. His main research interests include digital signal processing, computer vision, pattern recognition, and virtual reality.



Gaoyun AN received his B.Sc. degree in biological engineering and Ph.D. degree in signal and information processing from Beijing Jiaotong University, Beijing, China, in 2003 and 2008, respectively. He is currently Associate Professor with the Institute of Information Science, Beijing Jiaotong University, Beijing, China. His main research interests include image processing, computer vision, and pattern recognition.



Yi JIN received her Ph.D. degree in signal and information processing from the Institute of Information Science, Beijing Jiaotong University, Beijing, P.R. China, in 2010. She is currently Associate Professor in the School of Computer Science and Information Technology, Beijing Jiaotong University.