Accumulating weighted segmentation in 3D face recognition

Quan Ju*, Haitao Hu and Yingfeng Wang

School of Computer and Information Engineering,
Henan University of Economics and Law,
Zhengzhou, Henan Province, China
Email: 943161083@qq.com
Email: 331619534@qq.com
Email: 358703464@qq.com
*Corresponding author

Abstract: In this paper, an accumulating weighted face segmentation approach based on the rigid level of human facial areas is introduced. A mass of 3D face data is measured and analysed to define the most expression-invariant region. Different locations or regions on the human face are observed to have dissimilar invariant levels. Thus, an accumulating weight method is proposed to represent the rigid degree under expression variations. In face identification experiments, performance by employing the accumulating weight is demonstrated to be higher than methods using the expression-invariant region and the full face, respectively. This accumulating weighted face segmentation approach outperforms other state-of-the-art methods in 3D face recognition experiments.

Keywords: face segmentation; expression variation; accumulating weight; face recognition in 3D.

Reference to this paper should be made as follows: Ju, Q., Hu, H. and Wang, Y. (2022) 'Accumulating weighted segmentation in 3D face recognition', *Int. J. Wireless and Mobile Computing*, Vol. 22, No. 1, pp.84–92.

Biographical notes: Quan Ju received his BEng degree from the South China University of Technology and PhD degree in Computer Science from the Department of Computer Science, the University of York, UK. Currently, he is working as an Associate Director of Information Management and Information System Department in the School of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou, China. His research interests include pattern recognition, image processing and data mining.

Haitao Hu received his PhD degree in Electromagnetic Field and Microwave Technology from Beijing University of Posts and Telecommunications, Beijing, China in 2010. Currently, he is an Associate Professor in the School of Computer and Information Engineering at Henan University of Economics and Law, Zhengzhou, Henan Province, China. His research fields include pattern recognition, image processing and digital watermarking.

Yingfeng Wang received her PhD degree in Computer Application Technology from Xidian University, Shaanxi, China in 2011. Now, she is an Associate Professor in the College of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou, Henan, China. Her research interests include embedded systems, multi-core architecture, parallel processing and high performance computing.

This paper is a revised and expanded version of a paper entitled 'Accumulating weighted segmentation in 3D face recognition' presented at the '4th Asian Conference on Artificial Intelligence Technology', Chongqing, China, 2020.

1 Introduction

In recent years, face recognition has become an active field of pattern recognition and bioinformatics. Although the two-dimensional face recognition technology is gradually mature and widely used, face recognition in 2D has some difficulties that are not easy to overcome, such as the problem of light condition variations. Since 3D face images or models can provide richer information than 2D face images, more and more researchers are turning their attention to the research of 3D face recognition. In the application of 3D face recognition in real scenes, the problem of facial expression variations, which can lead to a decrease in face recognition performance, is a key difficulty that cannot be ignored.

Considerable research efforts have been devoted to handle expression problems by treating a human face as a rigid subject, because some facial regions always remain unchanged even under expression variations (Erdogmus et al., 2012; Bornak et al., 2010; Miao and Krim, 2011). Other approaches employ the deformation algorithms to recover the expression distortions by extracting local features (Bowyer et al., 2006; Samir et al., 2006; Li et al., 2011; Smeets et al., 2013; Faltemier et al., 2008 employ the ICP algorithm to match the probe image and the gallery image, they chose 28 regions from 38 facial areas as the best combination to achieve maximum matching performance. Tang et al. (2013) proposed a local binary model using different numbers of sparse areas and dense areas in 3D face recognition. Lei et al. (2014) extracted angular radial signature 3D facial feature from semi-rigid facial regions to improve the discriminating ability, then match faces by using kernel PCA and SVM. It is noteworthy from the above discussion that localising regions least affected by expressions is the key task for 3D face recognition.

Another practical scene that often appears in face recognition is that the face image may be blocked by masks, glasses, etc. Although numerous 3D face recognition approaches focus on handling the expression variations, there are relatively few researchers paying their attention on the occlusion problem in 3D face recognition (Zhou and Xiao, 2018). Colombo et al. (2006) presented a restoration strategy for occluded 3D face surface. Occluded regions are first localised according to their effects on the projections of faces in a suitable face space. These occluded regions are also can be recovered by using normal regions which are not occluded. Drira et al. (2013) presented a 3D statistical shape analysis method to deal with expression/pose variations and other occlusions. They remove the occluding object and create a radial shape statistical model to repair incomplete curves. Alyuz et al. (2013) introduced a robust and efficient method to perform face recognition in 3D even with a large amount of occlusions by dividing the face into many significant elements such as chin, mouth, cheek, eye etc. Differences between faces based on individual region matching are then combined to generate the overall similarity score. This composite scheme can automatically detect occlusions and then deal with them. As can be seen from approaches and methods mentioned above to overcome the occlusion problem, in addition to trying to use existing data to repair the missing data, one possible method is to localise and use the partial data of each key position on the remaining face image for face recognition.

Face segmentation or region localisation are very useful and important to overcome expression variations and occlusions problems in 3D face recognition. Local features extracted from different sub-regions have different abilities to handle these problems. For example, some researchers point out that the nose region is the most rigid area and does not change its shape under expression variations (Faltemier et al., 2008; Queirolo et al., 2010; Xue et al., 2009). However, they do not precisely tell the relationship between the distance from the nose tip and the degree of invariance. In this paper, a novel 3D face segmentation approach is introduced. The expression variation level at different positions is measured by analysing the degree of invariance under facial expression variations. Regions least affected by expression variations can be localised and defined. Weight values are also accumulated and assigned to all positions or sub-regions according to the rigid degree to expression variations. Combined with a face recognition algorithm based on a 3D surface descriptor, in the face recognition experiments conducted on the FRGC database which contains expression variations and data loss/occlusion, the results show that this method outperforms similar face segmentation methods.

In brief, our contributions are: (1) Through the analysis of the 3D face database, the distribution and location of facial expression invariant regions are discovered and visualised. (2) By segmenting the face into different regions, and assigning corresponding weight according to the invariance of expressions, a good foundation is established for the subsequent face recognition tasks. (3) Experimental results on FRGC database which is at present the most complex and largest 3D face database demonstrate that the proposed segmentation method is a simple, easy to apply and extremely effective approach to handle the problem of facial expression variations.

2 Related work

Dealing with facial expression variations is an urgent need but a major challenge in 3D face recognition. Recently, some researchers try to apply deep network to train 3D face images, so as to obtain promising performance in face recognition without implementing face segmentation and region positioning, even under facial expression variations (Gilani and Mian, 2018). Zheng et al. (2019) proposed an end-to-end face recognition system combining the geometric invariants, histogram of oriented gradients and the fine-tuned ResNet. Zhang et al. (2019) used a PointNet++ like network to extract face feature directly from point clouds. These methods usually require the design of a complex Deep Network or CNN, and also require laborious fine tune to obtain satisfying experimental results.

The instinctive problem-solving direction is to find these facial areas which are not or less influent by expression variations. Lim et al. (2018) proposed a 3D face component analysis technique to segment face by deformation face models to a fit template model. Meaningful facial regions can be found depending on the corresponding part of the template model. Kakadiaris et al. (2007) designed an Annotated Face Model (AFM) to handle the expression problems. Faces are fitted to the AFM model by a combination of Spin image, ICP and Simulated Annealing (SA) alignment algorithms. Then aligned models produce deformatting images to implement match by measuring the wavelet transformation distance metric. As the model fitting inevitably has considerable errors, almost all key facial areas cannot be located with acceptable accuracy in their experiments. Faltemier et al. (2008) used a Region Ensemble for FacE Recognition (REFER) algorithm to exploit face sub-regions that are relatively invariant under expression variations. Matchings between corresponding sub-regions are combined and processed by a committee of classifiers to improve face recognition performance. Li et al. (2014) used a learning-based strategy to determine the quantification weights of local facial patches according to the discriminating ability in face recognition under expression variations. Results in their experiments show that quantification weights learned for local patch in 3D are dramatically different from those of 2D faces. Sghaier et al. (2018) presented an automatic technique by using the anthropometric proportions and the cropping filter to extract the area of interest which is unaffected by facial expression. They use the highest z-value as a nose tip of the 3D face and use the concavity part to detect the eye corners. However, in a database with a large angle of pose variation, such an assumption may not be 100% true.

Some researchers suggest that upper facial areas such as nose and forehead are rigid facial locations which are least affected by expression changing. These areas can be marked as expression invariant regions (Lu et al., 2006; Ju and Wang, 2012). Other facial areas such as mouth, cheek and locations around eyes produce surface changes according to what kind of expressions are generated. Defining or localising expression-invariant regions actually require precise detection of these facial features. Unfortunately, most facial features such as eyes and mouth are quite difficult to be localised and detected accurately enough by utilising current techniques (Romero and Pears, 2008; Segundo et al., 2007)). Localising of some important facial features such as the forehead even depends on the detection of eyes. The nose area is probably the only exception. The nose region is the most stable area under expression variations, because this region has only one Facial Action Unit (FAU) which is closely related to facial expressions (Hager et al., 2002; Ekman and Friesen, 1978). All anatomical facial expressions are generated by actions of one or more FAUs. Table 1 shows the number of FAU in different facial locations. It is very possible to accurately detect and locate the nose tip.

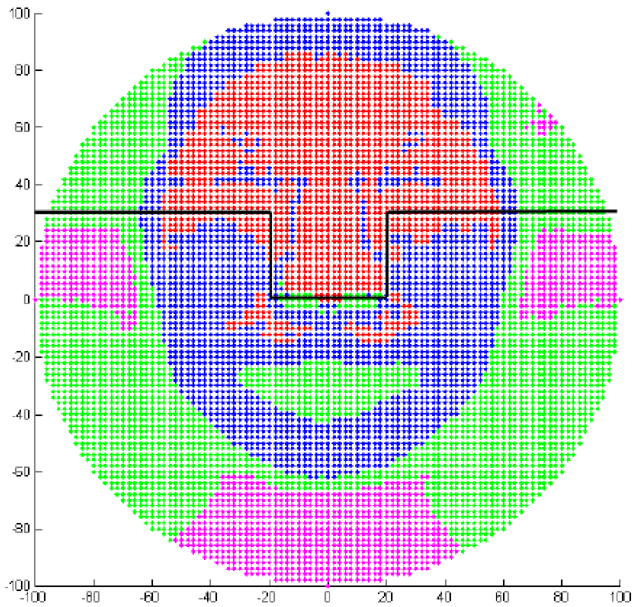
Table 1 The number of FAU in different facial regions

Region	*Number of FAU*
Nose area	one
Forehead	two or more
Eyes	five or more
Checks	five or more
Lower face area/mouth	20 or more

3 Accumulating weighted face segmentation

Some researchers have already successfully identified and localised the nose tip in 3D face (Faltemier et al., 2008; Queirolo et al., 2010; Romero and Pears, 2008; Ju et al., 2009; Wang et al., 2010; Mian et al., 2007). A face surface in 3D is often represented by a group of 3D points. The basic face surface could be cropped from the original face data by defining the nose tip as the sphere centre. Cropped faces can be aligned by using the Iterative Closest Point (ICP) algorithm (Ju, 2013). When all points in 3D are projected to the *xoy* plane, the nose tip on *xoy* plane is the origin of the projection circle. According to approaches mentioned above to detect/crop and align face in 3D, especially in FRGC data set, pose variations can be completely corrected and only expression variations left in 3D faces. Consequently, face surface changes in 3D caused by expression variations can be measured. Expression variant levels for each sampling positions is then revealed by calculating within-class *z* differences values for every individual in 3D face database. Based on the ground truth data in the FRGC database, Figure 1 can be generated to show different expression variant degrees at each sampling position by calculating Root Mean Square Error (RMSE) of these values for each individual in data set.

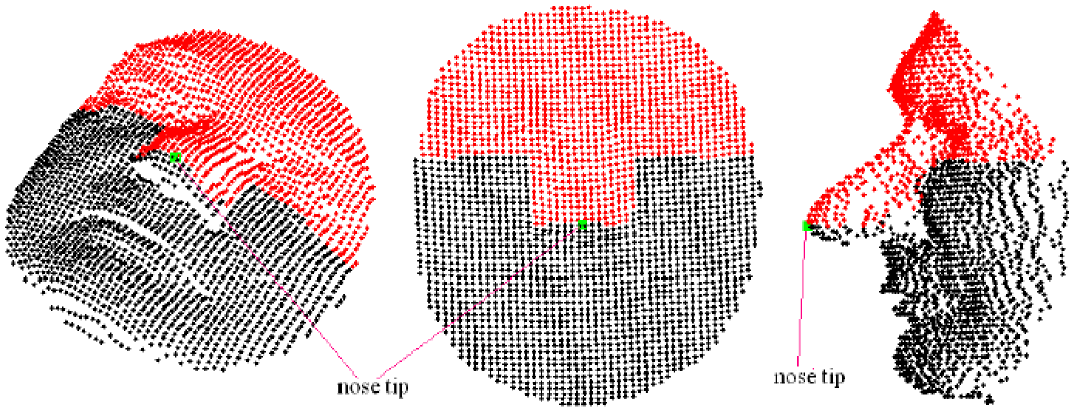
Figure 1 The upper region has lowest RMSE values $(<1.5\,\text{mm})$



The black lines shown in Figure 1 are possible borders of expression invariant regions. The quadrangular nose area and the area above the eyes shown in Figure 2 can be defined as the most expression invariant region. The red region in Figure 2 remains relatively unchanged even complicated facial expressions are produced. As a result, compared to other facial areas, this region will have more influence in face recognition.

As discussed in Section 2, FAU is related to facial expression. The nose region only has one FAU, thus it has the highest tolerance to expression variations. Some face recognition approaches indicate that even in the expression invariant region defined above, different positions also have different expression invariant abilities (Faltemier et al., 2008; Queirolo et al., 2010). Therefore, these different expression-invariant levels should be considered in face recognition tasks. The nose region is granted the highest weight value during implementing face recognition. Other regions have different influence or weight in face recognition.

Figure 2 The upper area is the expression invariant region



In some approaches such as Faltemier et al. (2008) and Mian et al. (2007), it is worth noting that the number of times that some regions participate in the face recognition calculation is different. The position/region around the nose tip has been used the most times in face recognition. The number of times that a face area/position participates in the face recognition calculation decreases as its distance moves away from the nose tip.

In this paper, based on results shown in Figure 1, we divide the whole face into two parts shown in Figure 2 the red upper face area is regarded as expression invariant region. However, every position has different weight according to its distance to the expression invariant region. In order to reduce computational complexity and facilitate subsequent face matching task, several radii can be defined to represent the differences of distance as shown in Figure 3.

By granting different weight values to different regions, a weight vector w is introduced. Because the position of the lower face including the mouth will be severely affected by changes in facial expressions, this area only obtains the basic weight value '1' in face recognition. Although the outermost circle of the upper face area is located in the area where the expression is less affected, it will only get a slightly higher weight '2' due to the influence of hair noise. As the diameter of the half ring areas in the upper face gradually decreases, the corresponding region weight increases. The rectangular area around the nose tip gets the highest weight value '9'. As a result, each 3D point of a face image will obtain a weight value and a weight vector w of this face image is then generated. The relationship between weight values and the position is shown in Figure 4.

Figure 3 The radii in upper expression invariant region are defined as 10 mm, 20 mm, 30 mm, 40 mm, 50 mm, 60 mm, 70 mm

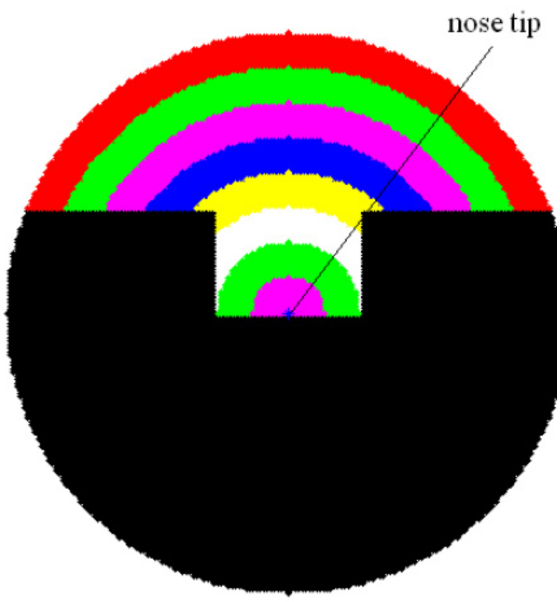
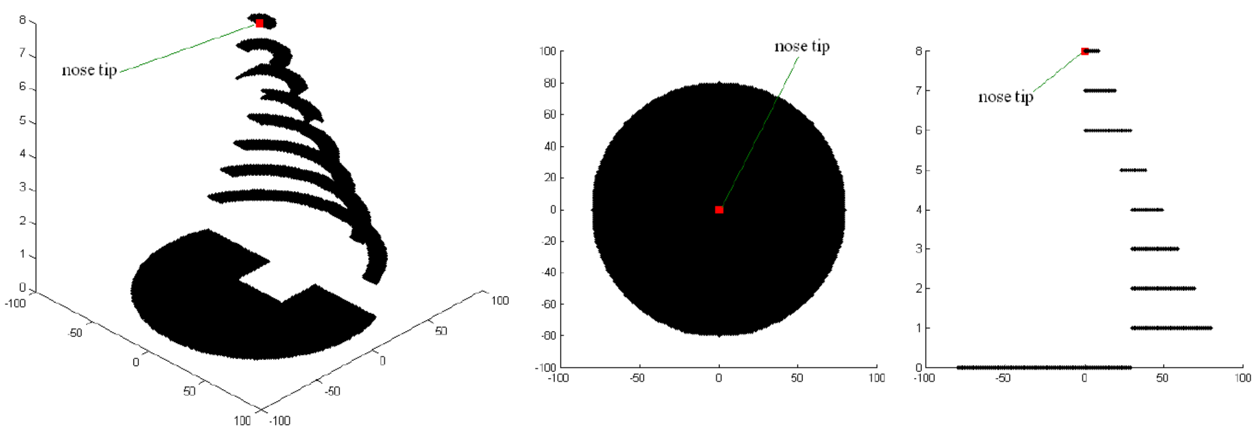


Figure 4 Regions are granted different weight based on their distances to nose tip, which is the most expression-invariant location



4 Evaluation by implementing face recognition

4.1 3D face database

In this paper, a face identification experiment is designed to evaluate the accumulating weighted segmentation. 4007 faces of 466 individuals in 3D from FRGC v2 database are selected as the experimental data (Phillips et al., 2005). Each individual has a number of faces with several different expressions. Some faces show very low data quality and some faces also have distortion because of poor standing still in data acquisition. And some faces lose data around their noses for some reason. Occlusions near mouth, eyes and forehead also can be observed in these faces. Defective and incomplete faces are shown in shown in Figure 5. On the other hand, to a certain extent, these data quality flaws make the FRGC v2 database a suitable data set to evaluate face segmentation method under expression variations and occlusions by implementing face recognition.

In this paper, three face identification experiments are defined by choosing the whole face area, the expression-invariant region defined above and the accumulating weight method respectively. There are 248 neutral faces (248 individuals) selected as the gallery data set. The rest 3759 faces with several expressions in FRGC v2 database constitute the query set. Matching neutral faces with expressive faces can evaluate the abilities of different segmentation methods to resist expression variations in face recognition.

4.2 Face matching algorithm

Inspired by some approaches by Grimson and Lozano-Perez (1984); Stein and Medioni (1991); Chua and Jarvis (1997); Xu et al. (2006) and Ankerst et al. (1999) , a 3D Multi shell shape descriptor is utilised to represent the surface around a point. A point P with its neighbouring P_i together can form a 3D shape which is a 1-ring mesh shown in Figure 6. Angles between all $P_i - P$ and the vertex normal N_p can be computed:

$$\theta_i = \arccos\left(\frac{(P_i - P)\cdot N_p}{|P_i - P||N_p|}\right) \tag{1}$$

Every point P has its neighbouring point set $n(P) = P_1, P_2, ..., P_n$ and one angle set $\theta(P) = \theta_1, \theta_2, ..., \theta_n$ (n is the number of neighbouring points). An indicator can show the convex or concave level of this 1-ring mesh by calculating the mean of all $\theta(P)$. For example, if the mean θ is greater than 90°, the mesh can be considered as a convex shape, and vice-versa. However using mean θ may not present enough details of the subtlety curvature shape in 3D. Standard Deviation (STD) of θ can be the reinforcement to describe more details of the 3D shape. The sphere also can be separated into more shells to generate a multi shell surface descriptor as shown in Figure 7.

STD and mean of θ are computed for every shell to generate the shape descriptor: $[std_1, std_2, ..., std_n]$, $[mean_1, mean_2, ..., mean_n]$, which denotes a surface around a point P. In FRGC face database, a 3D face is represented by a group of points, thus a face can be transferred to a group of shape descriptors. Shape descriptors of a 3D face can be described by an $m \times n$ matrix (m represents the number of shells, n signifies the number of points). A shape descriptor represents the surface around a particular point. In the matching between two faces, points selected from each face have been matched in pairs. If two points have similar shape descriptors, then two shapes around these two points can be considered the same. The different between two shape descriptors is the similarity degree of these two points. Faces belong to the same individual should have similar shape in 3D especially in expression invariant areas (Faltemier et al., 2008; Queirolo et al., 2010; Mian et al., 2007). The number of matching points which is considered as a correct match can be the score of similarity between two faces. In this paper, based on point-matching a face recognition method is employed using the shape descriptor introduced above to estimate the similarity of two faces. Results of 3D face identification experiments are used to evaluate this segmentation method.

Figure 5 Examples of faces having quality flaws

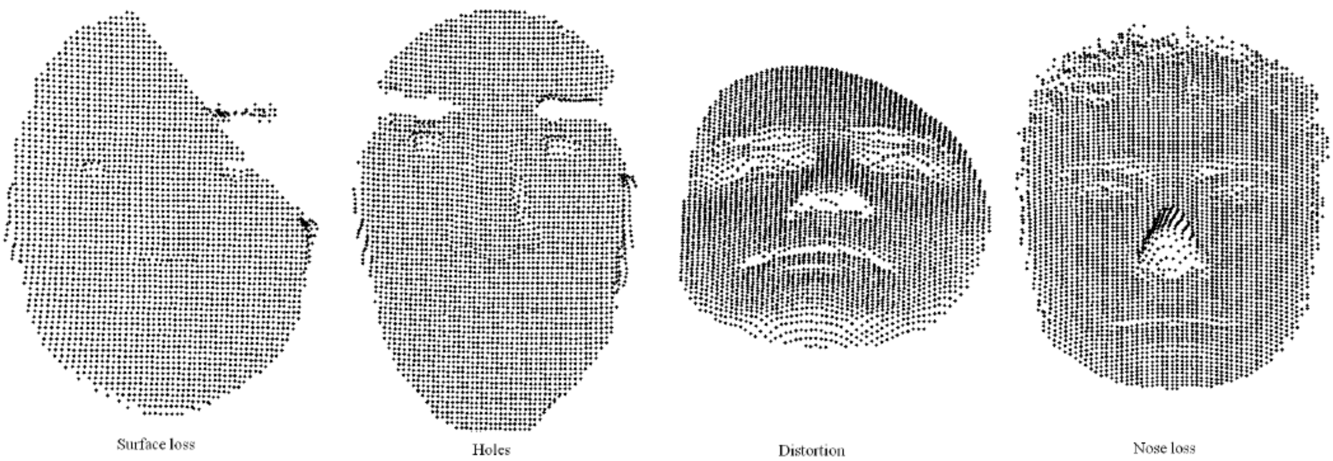


Figure 6 The 1-ring mesh created by *P* and its neighbouring point

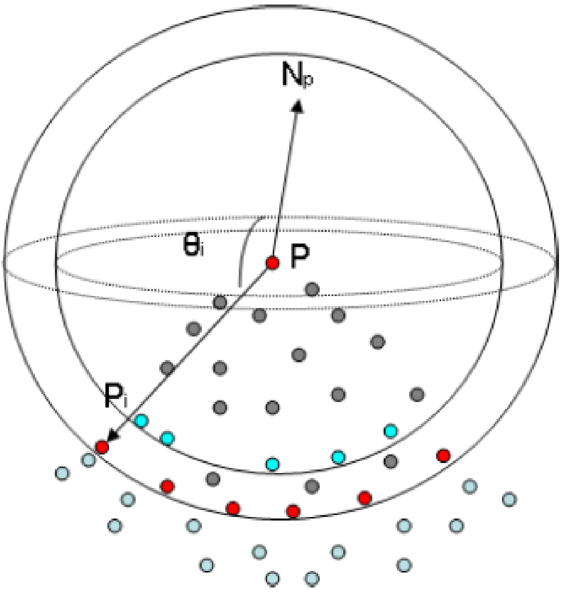
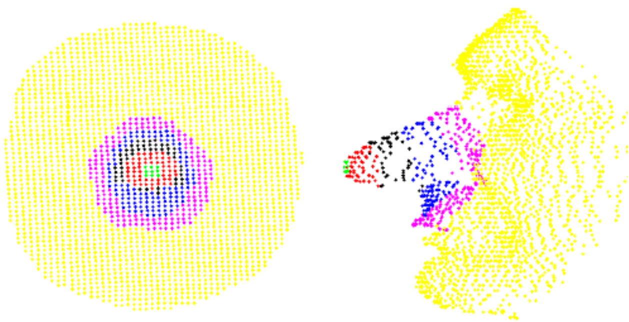


Figure 7 A 3D surface multi shell surface descriptor generated by several shells around a point



Five shells are designed in a shape descriptor (radii: $5\text{mm}, 10\text{mm}, 15\text{mm}, 20\text{mm}, 25\text{mm}$). The distance of two shape descriptors can be described by two distance vectors – dm and ds. Two pieces of 3D surface will be considered that they share identical shapes, if differences in matching distance vectors are both lower than particular thresholds ε_m and ε_s, respectively.

A set of steps to match two faces is shown in Algorithm 1. In steps 4–5, the STD θ vector and the mean θ vector shape descriptor vectors are calculated for the query point and the gallery point respectively. In steps 6–7, distances of STD and Mean vector between two points are computed. In steps 9–19, in order to reduce the effect of noise, thresholds ε_m and ε_s are defined to filter dm and ds. ndm represents the number of values in dm below threshold ε_m and nds is the number of ds below threshold ε_s. Another two discrimination thresholds t_m and t_s are designed to adjust noise tolerance ability and the power to distinguish faces from different persons in steps 20–21.

Algorithm 1

Require: A query face Q and a gallery face G

1: $S = 0$
2: **for** each point $p \in Q$ **do**
3: calculate vectors: m and s of the shape descriptor at the point p;
4: select a point p' in face G in order;
5: calculate m' and s' at the point p';
6: $dm = |m - m'|$;
7: $ds = |s - s'|$;
8: $ndm = 0$
9: **for** $dm_i \in dm$ **do**
10: **if** $dm_i < \varepsilon_m$ **then**
11: $ndm{+}{+}$;
12: **end if**
13: **end for**
14: $nds = 0$;
15: **for** $ds_i \in ds$ **do**
16: **if** $ds_i < \varepsilon_s$ **then**
17: $nds{+}{+}$;
18: **end if**
19: **end for**
20: **if** $ndm \geqslant t_m$ and $nds \geqslant t_s$ **then**
21: $S{+}{+}$;
22: **end if**
23: **end for**

Thresholds are carefully optimised to tolerate noises and keep the accuracy in the meantime. A face recognition experiment on the FRGC v1 database can be very helpful to determine these thresholds. If '1' represents a correct match and '0' means dissimilarity in matching, the result of two face matching will be the sum of all 'point vs. point' matchings. The overall similarity score S can be defined as the following equation:

$$S = \sum_{i=1}^{N} (\mathcal{M}_i) \tag{2}$$

where *N* denotes number of points of the query face.

Unlike other methods that need to match many different face regions that overlap each other, our method directly applies weights. Multiple matching of overlapping regions means a waste and useless calculation. Our method grants regions or positions corresponding weight values, which is easy to implement in the similarity calculation of face matching. When weight values of different positions are applied on similarity calculation, the similarity score with weights can be computed as an upgrade of equation (2):

$$S_w = \frac{\sum_{i=1}^{N}\left(w_i \cdot \mathcal{M}_i\right)}{N} \tag{3}$$

where w is the weight value, S_w is the similarity score with weights and M_i represents the result of point-point matching.

4.3 Experiment results

Evaluation experiments are conducted on FRGC v2 database to simulate the performance of a real face identification system, because FRGC v2 is the largest and most difficult 3D face database at present. The experimental gallery data set consists of 248 neutral face images for 248 individuals, and the query data set includes the other 3759 face images with expression changes and some data occlusions/loss. All faces in query data set match every face image in the gallery data set. The face image with the highest similarity score (rank-one) is regarded as the identification result, and the rank-one face identification rate is then calculated.

Three different strategies to match face regions including the whole face, expression-invariant regions and the full face with accumulating weights are used to explore the advantage of the accumulating weighted method. Using full face area only, A 93.7% face identification rate is achieved. A slightly higher result is observed in the face identification using expression-invariant region. When the accumulating weight is applied in face matching, the identification accuracy is obviously improved as shown in Table 2. Experimental results demonstrate that using the accumulating weighted segmentation method can improve the performance of face recognition even under expression variations.

Table 2 Identification accuracies by employing different regions matching or segmentation methods

Region used in matching:	*Identification rate*
Whole face	93.7%
Expression invariant areas	94.2%
Accumulating weights applied	97.6%

Mian et al. (2007) and Faltemier et al. (2008) presented two similar segmentation methods. However, in the task of face recognition, both of them conduct matchings between corresponding regions respectively, and then merge the results together. Unlike these methods, our approach does not need to repeatedly compare/match many face regions that actually overlap each other, but directly applies accumulated weights in the similarity calculation. Unlike these two method trying to localise the expression-invariant regions, Li et al. (2015) quantify the relative discriminating ability for each facial patch/area to formulate or learn a weighted sparse representation model. Theoretically, poor data/image quality in training data set such as occlusion/data loss will lower the performance of their method. Other approaches based on deep network/deep learn such as Zhang et al. (2019); Zheng et al. (2019) and Gilani et al. (2018) are very popular in recently. Thanks to the advantages of deep network technique, these methods achieve promising results in face recognition. Comparing with these methods, we can show the rationality of our method to distribute weights for each region in face recognition. Table 3 shows that our segmentation method applying accumulating weights exceeds the accuracy of these state-of-the-art approaches in the same rank-one face identification experiment based on FRGC database.

Table 3 Compared with other state-of-the-art methods

Method	*Identification rate*
Mian et al. (2007)	96.2%
Faltemier et al. (2008)	97.2%
Li et al. (2015)	96.7%
Zhang et al. (2019)	92.7% without fine tuning
Zheng et al. (2019)	97.0% without fine tuning
Gilani et al. (2018)	97.1% without fine tuning
Ours	97.6%

5 Conclusion

In this paper, a novel accumulating weighted face segmentation method according to the expression variant degree is introduced. Firstly, expression-invariant regions are localised and defined by analysing 3D face data set. Then, various locations within expression-invariant regions are assigned different weight values. These weight values can be accumulated to represent the rigid degree under expression variations. Compared to methods using expression-invariant regions and the whole face respectively, the performance of face identification experiments by applying these accumulating weights is significantly improved. It shows that for face recognition, this method is more reasonable for face segmentation and region weight distribution. In addition, compared with the rank-one face identification results of some of state-of-the-art approaches, our method also has obvious advantages.

Acknowledgement

This paper is funded by Henan Provincial (China) Key Scientific Research Project for College and University Grant (No. 18B520004).

References

Alyuz, N., Gokberk, B. and Akarun, L. (2013) '3D face recognition under occlusion using masked projection', *IEEE Transactions on Information Forensics and Security*, Vol. 8, pp.789–802.

Ankerst, M., Kastenmuller, G., Kiegel, H. and Seidl, T. (1999) '3d shape histograms for similarity search and classification in spatial databases', *Proceedings of the 6th International Symposium on Spatial Databases*, pp.207–226.

Bornak, B., Rafiei, S., Sarikhani, A. and Abolghasem, B. (2010) '3D face recognition by used region-based with facial expression variation', *Proceedings of the 2nd International Conference on Signal Processing Systems*, IEEE, China.

Bowyer, W., Chang, K. and Flynn, P. (2006) 'A survey of approaches and challenges in 3D and multi-modal 3D+2D face recognition', *Computer Vision and Image Understanding*, Vol. 101, No. 1, pp.1–15.

Chua, C. and Jarvis, R. (1997) 'Point signature: a new representation for 3d object recognition', *International Journal of Computer Vision*, Vol. 25, No. 1, pp.63–85.

Colombo, A., Cusano, C. and Schettni, R. (2006) 'Detection and restoration of occlusions for 3D face recognition', *Proceedings of the IEEE International Conference on Multimedia and Expo*, Toronto, Ontario, Canada.

Drira, H., Amor, B., Srivastava, A., Daoudi, M. and Slama, R. (2013) '3D face recognition under expressions, occlusions, and pose variations', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 9, pp.2270–2283.

Ekman, P. and Friesen, W. (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, Palo Alto.

Erdogmus, N., Daniel, L. and Dugelay, J. (2012) 'Probabilistic fusion of regional scores in 3D face recognition', *Proceedings of the International Conference on Image Processing*, pp.1–5.

Faltemier, T., Bowyer, K. and Flynn, P. (2008) 'A region ensemble for 3d face recognition', *IEEE Transactions on Information Forensics and Security*, Vol. 3, No. 1, pp.62–73.

Faltemier, T., Bowyer, K. and Flynn, P. (2008) *Using Multi-Instance Enrollment to Improve Performance of 3D Face Recognition*, Elsevier Science Inc.

Gilani, S.Z. and Mian, A. (2018) 'Learning from millions of 3D scans for large-scale 3D face recognition', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1896–1905.

Grimson, W and Lozano-Perez, T. (1984) 'Model-based recognition and localization from tactile data', *IEEE International Conference on Robotics*, Atlanta, GA.

Hager, J., Ekman, P. and Friesen, W. (2002) *Facial Action Coding System*, A Human Face, Salt Lake City, UT.

Ju, Q. (2013) 'Robust binary neural networks based 3D face detection and accurate face registration', *International Journal of Computational Intelligence Systems*, Vol. 6, pp.669–683.

Ju, Q. and Wang, Y. (2012) 'Accurate 3D nose tip localization using orientation invariant surface descriptor', *International Journal of Digital Content Technology and its Applications*, Vol. 6, pp.144–152.

Ju, Q., O'Keefe, S. and Austin, J. (2009) 'Binary neural network based 3D facial feature localization', *Proceedings of the 2009 international joint conference on Neural Networks*, pp.843–850.

Kakadiaris, I., Passalis, G., Toderici, G., Murtuza, N. and Theoharis, T. (2007) 'Three-dimensional face recognition in the presence of facial expression: an annotated deformable model approach', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 4, pp.671–680.

Lei, Y., Bennamoun, M. and Hayat, M. (2014) 'An efficient 3D face recognition approach using local geometrical signatures', *Pattern Recognition*, Vol. 47, No. 2, pp.509–524.

Li, H., Huang, D. and Lemaire, P. (2011) 'Expression robust 3D face recognition via mesh-based histograms of multiple order surface differential quantities', *IEEE International Conference on Image Processing*.

Li, H., Huang, D., Morvan, J-M., Chen, L. and Wang, Y. (2014) 'Expression-robust 3D face recognition via weighted sparse representation of multi-scale and multi-component local normal patterns', *Neurocomputing*, Vol. 133, pp.179–193.

Li, H., Huang, D., Morvan, J-M., Wang, Y. and Chen, L. (2015) 'Towards 3D face recognition in the real: a registration-free approach using fine-grained matching of 3D keypoint descriptors', *International Journal of Computer Vision*, Vol. 113, pp.128–142.

Lim, S.J., Hwang, B.W., Yoon, S.U., Jin, S.C. and Park, C.J. (2018) 'Automatic 3D face component analysis technique', *IEEE International Conference on Consumer Electronics (ICCE)*, IEEE, USA.

Lu, X., Jain, A. and Colbry, D. (2006) 'Matching 2.5D face scans to 3D models', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 1, pp.31–43.

Mian, A., Bennanmoun, M. and Owens, R. (2007) 'An efficient multimodal 2D-3D hybrid approach to automatic face recognition', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 11, pp.1927–1943.

Miao, S. and Krim, H. (2011) 'Robustness and expression independence in 3d face recognition', *IEEE Workshop on Signal Processing Systems, Proceedings*, pp.289–292.

Phillips, P., Flynn, P., Scruggs, T and Bowyer, K. (2005) 'Overview of the face recognition grand challenge', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.947–954.

Queirolo, C., Silva, L., Bellon, O. and Segundo, M. (2010) '3D face recognition using simulated annealing and the surface interpenetration measure', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 2, pp.206–219.

Romero, M. and Pears, N. (2008) '3D facial landmark localisation by matching simple desciptors', *Proceedings of the 2nd IEEE Int. Conf. Biometrics: Theory, Applications and Systems*.

Samir, C., Srivastava, A. and Daoudi, M. (2006) 'Three-dimensional face recognition using shapes of facial curves', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp.1858–1863.

Segundo, M., Queirolo, C., Bellon, O. and Silva, L. (2007) 'Automatic 3D facial segmentation and landmark detection', *Image Analysis and Processing*, pp.431–436.

Sghaier, S., Farhat, W. and AbSouani, C. (2018) 'Novel technique for 3D face recognition using anthropometric methodology', *International Journal of Ambient Computing and Intelligence (IJACI)*, Vol. 9, pp.60–77.

Smeets, D, Keustermans, J., Vandermeulen, D. and Suetens, P. (2013) 'meshSIFT: local surface features for 3D face recognition under expression variations and partialdata', *Computer Vision and Image Understanding*, Vol. 117, No. 2, pp.158–169.

Stein F. and Medioni G. (1991) 'Structural hashing: Efficient three dimensional object recognition', *Proceedings of Computer Vision and Pattern Recognition*, pp.244-250.

Tang, H., Yin, B., Sun, Y. and Hu, Y. (2013) '3D face recognition using local binary patterns', *Signal Processing*, Vol. 93, No. 8, pp.2190–2198.

Wang, Y., Liu, J. and Tang, X. (2010) 'Robust 3D face recognition by local shape difference boosting', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, pp.1858–1870.

Xu, C., Li, S.Z., Tand, T. and Quan, L. (2009) 'Automatic 3D face recognition from depth and intensity Gabor features', *Pattern Recognition*, Vol. 42, No. 9, pp.1895–1905.

Xu, C., Tan, T., Wang, Y. and Quan, L. (2006) 'Combining local features for robust nose location in 3D facial data', *Pattern Recognition Letters*, Vol. 27, No. 13, pp.1487–1494.

Zhang, Z., Da, F., Khalid, F. and Yu, Y. (2019) 'Data-free point cloud network for 3D face recognition', *arXiv:1911.04731vi*.

Zheng, S., Rahmat, R.W.O., Khalid, F. and Nasharuddin, N.A. (2019) '3D texture-based face recognition system using fine-tuned deep residual networks', *Peer Journal of Computer Science*, Vol. 5, No. 6.

Zhou, S. and Xiao, S. (2018) '3D face recognition: a survey', *Human-centric Computing and Information Sciences*, Vol. 8, No. 35, pp.1–27.