

# 3D Face Recognition Using an Expression Insensitive Dynamic Mask

Sadegh Salahshoor and Karim Faez

E.E. Dept. Electrical Engineering  
Amirkabir University of Technology (Tehran Polytechnic)  
Tehran, Iran

{S.Salahshoor, KFaez}@AUT.ac.ir

**Abstract.** Human face recognition is one of the most popular biometric approaches. In last decade 3D face recognition attracted much attention. In this paper, we present an automatic face recognition algorithm and demonstrate its performance on the Bosphorus 3D face database. A novel Dynamic mask is used to segment automatically the regions of face which are less sensitive to expressions. We applied a multilayer perceptron (MLP) to compute maskable region (MR). MR shows which percentage of face image pixels must be masked to produce the expression insensitive binary mask for 3D faces. We applied a modified nearest neighbor classifier for identification. We only used one neutral frontal face of each subject as gallery images and tested our algorithm with emotional expression images. The identification rate obtained is 85.36 percent in non-neutral expression.

**Keywords:** Biometrics, 3D Face Recognition, Facial Expression.

## 1 Introduction

Face recognition is a great challenge for computer vision and is one of the most attractive biometric approaches. In last decade, great advantages occurred in face recognition, but the corresponding methods are not robust and reliable enough. [1]

The performance of face recognition systems can be affected by five key factors: 1) Illumination variations; 2) Pose changes; 3) Expression variation; 4) Time delay; 5) Occlusions. [1]

Recent researches on 2D and 3D face recognition systems show that 3D face recognition is more robust to illumination variations and pose changes. [2]

Chua et al [3] used iterative closest point (ICP) for 3D face recognition and they applied a Gaussian distribution to separate the rigid and non-rigid parts of the face. Zhong et al [4] used 3D depth images. Xu et al [10] applied the principal component analysis (PCA) to construct the 3D eigenfaces and used a nearest neighbor classifier to recognition. Queirolo et al [5] applied the simulated annealing (SA) for registration with surface interpenetration measure (SIM) as similarity measure. They combined SIM values of four different face regions: circular and elliptical areas around the nose,

forehead, and the entire face region to achieve recognition. Xu et al [6] applied Gabor wavelet filter on the depth and intensity images; they used linear discriminate analysis (LDA) and adaboost learning to extract the features. Mian et al [7] used the spherical face representation (SFR) and scale-invariant feature transform (SIFT) to reduce the number of candidate faces. Since the eye-forehead and nose regions are less sensitive to expressions, they used a modified ICP on these regions to achieve recognition.

In this paper, we propose an automatic 3D face recognition system robust to expression variations. (see fig. 1) we applied a multilayer perceptron (MLP) to compute the maskable region (MR) for each probe image. This value is used to produce a dynamic mask which separates the least sensitive region of 3D face Image for recognition.

The rest of this paper is organized as follows. Section 2 introduces the preprocessing procedure including 3D face denoising and normalization which are very important to robust face recognition. Section 3 describes the proposed Artificial Neural Network dynamic mask structures. Section 4 reports our experimental result. Finally, in section 5, a conclusion summarizes this paper.

## 2 Preprocessing

Since most of acquired 3D face samples are noisy and without alignment, a preprocessing approach is needed. Preprocessing includes three steps.

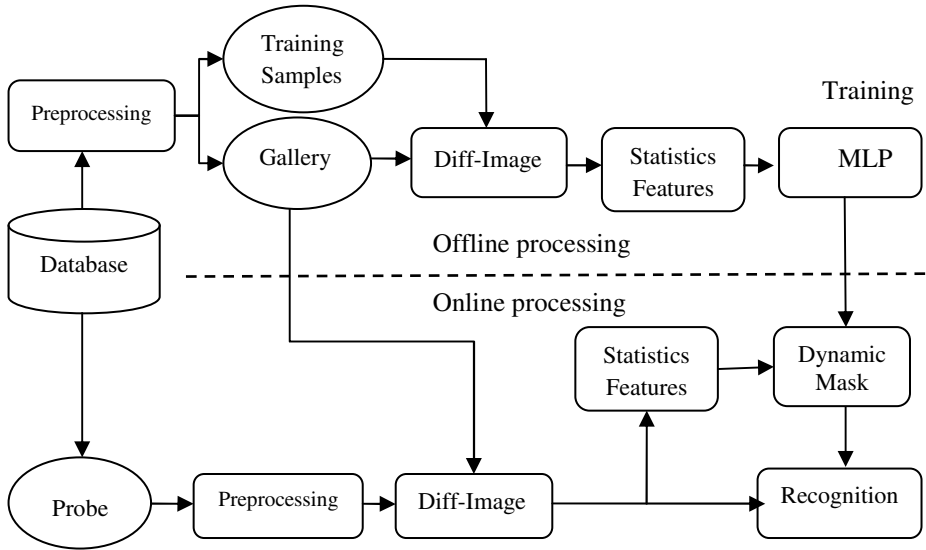
At the first step, 3D face images must be denoised, three Gaussian filters are applied to remove the spikes and reduce the noises. These filters smooth the data with different variations. But smoothing may cause to lose some sharp details of images, so it is a trade-off between the denoising and keeping the important information of images.

At the second step, the region of interest (ROI) in the face is selected. For this purpose, we detect the nose tip at first. The point with highest z value in the middle of face region is selected as approximate nose tip. Then spherical region around the nose tip is cropped from the denoising data.

At the last step, the 3D faces have been aligned to a standard posture by ICP algorithm [8]. ICP is used to align the forehead and nose regions of a 3D face (see fig. 2b), which is less sensitive to facial expression, with a reference image (see fig. 2a). The reference image is a mean image of the neutral 3D faces of database. Finally the aligned face is normalized to a 160x135 depth image (see fig. 2c). The block diagram of the preprocessing procedure is shown in fig. 3.

## 3 3D Face Recognition Using Dynamic Mask

The algorithm starts by finding the maskable region (MR) for each probe image. An Artificial Neural Network (ANN) is applied to calculate a list of MRs for each gallery images. The minimum size MR in the list is selected as desired MR. An MR shows which percentage of face image pixels must be omitted by making a binary mask.



**Fig. 1.** Block diagram of our recognition algorithm. The dashed line separates the online and offline phases.

### 3.1 Artificial Neural Network (ANN)

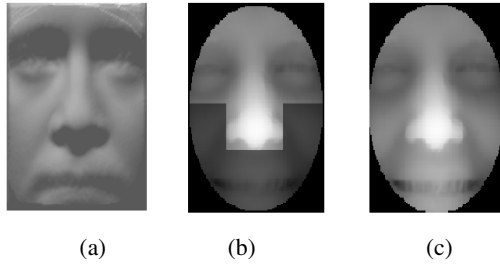
A Multilayer perceptron (MLP) is applied to compute MR. The MLP is a feed forward ANN which comprises of multiple layers of nodes. Each node is fully connected to nodes of the previous and next layers. The nodes in the same layer do not connect together. The nodes are called neurons and each of them has a nonlinear activation function. Hyperbolic tangent and logistic function are the most well-known activation functions which are sigmoid, and described by:

$$f(x) = \tanh(x). \quad (1)$$

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2)$$

where the former ranges from -1 to 1 and the latter ranges from 0 to 1.

Back propagation (BP) method is used to train an MLP. BP is a kind of supervised learning algorithm. In this algorithm, weights of the nodes are updated by back propagating the errors between desired and actual outputs.

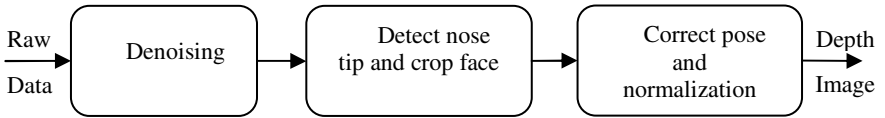


**Fig. 2.** (a) Reference image (mean face). (b) Forehead region which is used to align the expression faces. (c) A normalized depth image.

We used a simple three layers, which has 11 nodes in the input layer, nine neurons in the hidden layer and only one node in the output layer. Inputs are statistical features of the Differential-Image which is described by:

$$Diff_{Image} = |probe - gallery|. \tag{3}$$

where  $|x|$  denotes the absolute number of  $x$ 's. Then we reshaped this  $M \times N$  matrix to a  $NM \times 1$  Differential-vector and sorted it into ascending order.



**Fig. 3.** Block diagram of the preprocessing Procedure

The statistical inputs are minimum, maximum, median, 1<sup>st</sup> and 3<sup>rd</sup> quartile and 9<sup>th</sup> decile of Differential-vector. The others Inputs are:

a) arithmetic mean:

$$mean = \frac{1}{n} \sum_{i=1}^n x(i). \tag{4}$$

b) interquartile mean:

$$IQmean = \frac{2}{n} \sum_{i=\frac{n}{4}+1}^{\frac{3n}{4}} x(i). \tag{5}$$

where  $x$  is an ascend vector.

c) Geometric mean:

$$Gmean = (\prod_{i=1}^n x(i))^{\frac{1}{n}}. \tag{6}$$

d) harmonic mean:

$$Hmean = \frac{n}{\sum_{i=1}^n \frac{1}{x(i)}}. \quad (7)$$

and e) the standard deviation:

$$std = \sqrt{\frac{1}{n} \sum_{i=1}^n (x(i) - mean)^2}. \quad (8)$$

### 3.2 Dynamic Mask

Differential-Images (Eq. 3) are used to make binary masks for each pair of probe-gallery faces. These masks separate the regions of face which are the least affected by facial expression. A dynamic threshold for each pair of probe-gallery faces is selected to make the dynamic mask.

$$DTh = Diff_{vector}(N * (1 - MR)). \quad (9)$$

where  $DTh$  denotes the dynamic threshold,  $Diff_{vector}$  is a vector corresponding to the facial regions of differential- image, which is sorted into ascending order.  $N$  denotes the length of  $Diff_{vector}$  and  $MR$  is the maskable region.

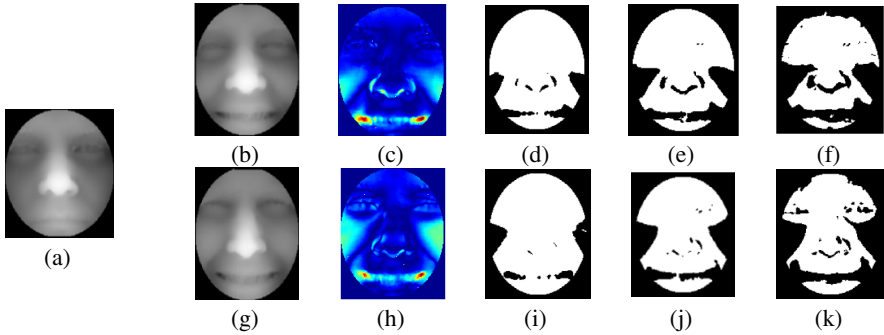
In order to make a dynamic mask, the pixels of differential-image whose magnitudes are less than the threshold are set to 1 and other pixels are set to zero. Some examples of dynamic mask are shown in fig. 4.

### 3.3 Classification

The classification is performed using a modified nearest neighbor algorithm (NN). NN computes the distances from a probe face to all gallery faces. The number of pixels equal to one in the dynamic masks, might be different for each pair of probe-gallery images; however, MRs are unique for each probe images. This can affect the accuracy of our algorithm, so we applied a modified mean squared error (MMSE) to measure the distance which is described by:

$$MMSE = \frac{1}{K} \sum_{i=1}^N \sum_{j=1}^M mask(i, j) * (P(i, j) - G(i, j))^2. \quad (9)$$

where  $P$  is a probe and  $G$  is a gallery face and the mask is a dynamic mask which is made for this pair of probe-gallery faces and all of them are  $N \times M$  matrices.  $K$  is the number of pixels in dynamic mask which are equal to unity.



**Fig. 4.** some examples of dynamic mask. (a) neutral expression. (b, g) happy expressions. (c, h) the differential Images. (d, i) dynamic masks (IP=0.1). (e, j) dynamic masks (IP=0.2). (f, k) dynamic masks (IP=0.3). first row (b-f) and (a) are belonging to the same person.

## 4 Experimental Result and Discussions

We tested our 3D face recognition method on the emotional expression faces of Bosphorus database [9].

### 4.1 Bosphorus 3D Face Database

Bosphorus 3D face database comprises 4666, 3D faces (shape and texture). It consists of 105 subjects (61 men and 44 women) in various expressions, pose and occlusion.

Bosphorus database comprises two set of expressions, the first set consists of the expressions which are known as action units (AUs) [10] and the second set comprise emotional expressions, which are anger, disgust, fear, happiness, sadness, and surprise, but all subjects do not have all the emotional expressions. In table 1, the frequency of each emotional expression in the database is given; also fig. 5 shows some 3D faces of Bosphorus database.

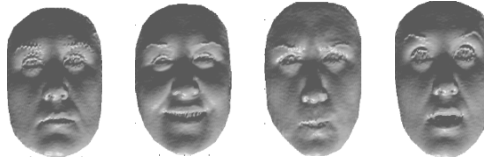
**Table 1.** Frequency of each of emotional expressions in Bosphorus database

anger	disgust	fear	happiness	sadness	surprise
71	69	70	105	66	70

### 4.2 Identification Result

In the experiments, the neutral expression face is used as the single gallery of a person, and the emotional expression faces are used as probes.

At offline processing approach, the best ignored rates (MR) for some samples of database are calculated, and then these samples are used to train the Neural Network. Since the Maskable Region range from 0 to 1, the logistic function (Eq.2) is used as activation function of MLP neurons. We used this trained MLP to make the dynamic mask at online processing.



**Fig. 5.** Some samples of Bosphorus 3D face database

We compare our proposed dynamic mask with three different face regions: circular and elliptical areas around the nose and the entire face, and applied some fundamental methods such as, 3D eigenfaces [11], and Gabor features for this comparison. Gabor filters of five scales and eight orientations were used to extract the Gabor features, and the PCA was applied to reduce the dimensionality of the feature space. In table 2, the comparing results are given. It shows that the proposed method achieve an average accuracy of 85.36 percent and the recognition range for different expressions range from 74.29 to 95.43 percent.

**Table 2.** Comparing rank-one identification rates of the proposed method with other methods

	Circular areas around the nose		Elliptical areas around the nose		Entire face		
	3D eigenface [11]	Gabor + PCA	3D eigenface [11]	Gabor + PCA	3D eigenface [11]	Gabor + PCA	The proposed method
anger	80.28	71.83	81.69	70.42	83.10	36.62	87.32
disgust	68.12	63.77	78.26	57.97	34.78	26.09	78.26
fear	80	72.86	72.86	68.57	35.71	32.87	92.86
happiness	58.10	75.24	54.29	59.05	12.38	23.81	74.29
sadness	84.85	83.33	81.82	80.30	84.85	53.03	95.45
surprise	82.86	75.71	64.29	71.43	24.28	32.86	91.43
<b>overall</b>	74.28	73.84	70.73	67.18	43.02	33.26	85.36

## 5 Conclusions

In this paper, we proposed a new 3D face recognition system uses an expression insensitive dynamic mask for identification. In this new proposed method, a MLP is applied to estimate the Maskable Region (MR) which is used to make the dynamic mask. This mask separates the less sensitive region of face, and finally a modified nearest neighbor classification algorithm is used to recognition. Experiments on the Bosphorus 3D face database show that our method is more effective than the region based method. On the Bosphorus database with one neutral face per individual in the gallery, the proposed method achieved the rank-one identification rate of 85.36 percent.

## References

1. Abate, A.F., Nappi, M., Riccio, D., Sabatino, G.: 2D and 3D Face Recognition: A survey. *Pattern Recognition Letters*, 1885–1906 (2007)
2. Bowyer, K., Chang, K., Flynn, P.: A Survey of approaches and challenges in 3D and multi-model 3D + 2D face recognition. In: *CVIU*, vol. 101, pp. 1–15 (2006)
3. Chau, C.S., Han, F., Ho, Y.K.: 3D human face recognition using point signature. *Automatic Face and Gesture Recognition*, 233–238 (2000)
4. Zhong, C., Sun, Z., Tan, T.: Robust 3D face recognition using learned visual codebook. *Pattern Recognition*, 1–6 (2007)
5. Queirolo, C.C., Silva, L., Bellon, O.R.P., Segundo, M.P.: 3D face recognition simulated annealing and the surface interpenetration measure. *IEEE Transaction, Pattern Analysis and Machine Intelligence* 32(2), 206–219 (2010)
6. Xu, C., Li, S., Tan, T., Quan, L.: Automatic 3D face recognition from depth and intensity Gabor features. *Pattern Recognition* 42, 1895–1905 (2009)
7. Mian, A.S., Bennamoun, M., Owens, R.: An efficient multimodal 2D-3D hybrid approach to automatic face recognition. *IEEE Transaction, Pattern Analysis and Machine Intelligence* 29(11), 1927–1943 (2007)
8. Besl, P.J., McKay, N.D.: A method for registration of 3D shapes. *IEEE Transaction, Pattern Analysis and Machine Intelligence* 14, 39–256 (1992)
9. Savran, A., Alyüz, N., Dibeklioglu, H., Çeliktutan, O., Gökberk, B., Sankur, B., Akarun, L.: Bosphorus Database for 3D Face Analysis. In: Schouten, B., Juul, N.C., Drygajlo, A., Tistarelli, M. (eds.) *BIOID 2008*. LNCS, vol. 5372, pp. 47–56. Springer, Heidelberg (2008)
10. Ekman, P., Friesen, W.V.: *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto (1978)
11. Xu, C., Wang, Y., Tan, T., Quan, L.: A new attempt to face recognition using 3D-eigenfaces. In: *The 6th Asian Conference on Computer Vision (ACCV)*, vol. 2, pp. 884–889 (2004)