

3D Face Recognition in Biometrics

CHAO LI, ARMANDO BARRETO
Electrical & Computer Engineering Department
Florida International University
10555 West Flagler ST. EAS 3970
33174 USA

Abstract: -Biometrics is the area of bioengineering that pursues the characterization of individuals in a population (e.g., a particular person) by means of something that the individual is or produces. Among the different modalities in biometrics, face recognition has been a focus in research for the last couple of decades because of its wide potential applications and its importance to meet the security needs of today's world. To date, most of the systems developed are based on 2D face recognition technology, which uses pictures for data processing. With the development of 3D imaging technology, 3D face recognition emerges as an alternative to overcome the difficulties inherent to 2D face recognition, i.e. sensitivity to illumination conditions and positions of a subject. But 3D face recognition still needs to tackle the problem of deformation of facial geometry that results from the expression changes of a subject. To deal with this issue, a 3D face recognition framework is proposed in this paper. It is composed of three subsystems: expression recognition system, expressional face recognition system and neutral face recognition system. A system for the recognition of faces with one type of expression (smile) and neutral faces was implemented and tested on a database of 30 subjects. The results proved the feasibility of this framework.

Key-Words: - face recognition, biometrics, 2D, 3D, range image, PCA, subspace, SVM, LDA

1 Introduction

Biometrics is a specific area of bioengineering. It pursues with recognition of a person through something the person has, i.e., the biological characteristics or something the person produces, i.e., behavioral characteristics. Examples of the former include finger print, iris, retina, palm print, face, DNA, etc. Examples of the latter include voice, handwriting, gait, signature, etc. Biometrics is used for identification or authentication in border control, e-commerce, ATM machine access control, crowd surveillance, etc. In recent years, biometrics gained more and more attention for its potential application in anti-terrorism.

Among the different modalities used in biometrics, the face is considered to be the most transparent one. It requires minimum cooperation from the subject. In some application scenario like crowd surveillance, face recognition probably is the only feasible modality to use. Face recognition is also the natural way used by our human beings in daily life. Therefore, face recognition has attracted many researchers from different disciplines, such as image processing, pattern recognition, computer vision, and neural networks. Face recognition scenarios can be classified into the following two:

- **Face verification** ("Am I who I say I am?") is a one-to-one match that compares a query face

image against a gallery face image whose identity is being claimed. (The gallery face images are the images which have been stored in the database.)

- **Face identification** ("Who am I?") is a one-to-many matching process that compares a query face image against all the gallery images in a face database to determine the identity of the query face. In the identification task, we assume that through some other methods we know that the person is in the database. The identification of the query image is done by locating the image in the database that has the highest similarity with the query image. In this paper, the face identification problem was addressed.

Most of the face recognition attempts that have been made until recently use 2D intensity images as the data format for processing. Varying levels of success have been achieved in 2D face recognition research. Detailed and comprehensive surveys can be found in [1, 2]. Although 2D face recognition has achieved considerable success, certain problems still exist. Because the 2D face images used not only depend on the face of a subject, but also depend on the imaging factors, such as the environmental illumination and the orientation of the subject. These two sources of variability in the face image often make the 2D face

recognition system fail. That is the reason why 3D face recognition is believed to have an advantage over 2D face recognition.

With the development of 3D imaging technology, more and more attention has been directed to 3D face recognition. In [3], Bowyer et al. provide a survey of 3D face recognition technology. Most of the 3D face recognition systems treat the 3D face surface as a rigid surface. But actually the face surface is deformed by different expressions of the subject. Therefore, systems that treat the face as a rigid surface are prone to fail when dealing with faces with expressions. The involvement of facial expression has become an important challenge in 3D face recognition systems. In this paper, we propose an approach to tackle the expression challenge in 3D face recognition. Because the deformation of the face surface is always associated with a specific expression, an integrated expression recognition and face recognition system is proposed. In section 2, a model of the relationship between expression and face recognition is introduced. Based on this model, the framework of integrated expression recognition and face recognition is proposed. Section 3 explains the acquisition of the experimental data used and the preprocessing performed. Section 4 outlines our approach to 3D face expression recognition. Section 5 explains the process used for 3D face recognition. Section 6 describes the experiments performed and the results obtained. Section 7 presents our discussion and conclusion.

2 Expression Recognition and Face Recognition

From the psychological point of view, it is still not known whether facial expression recognition information aids the recognition of faces by human beings. One of the experiments that support the existence of a connection between facial expression recognition and face recognition was reported in [4]. The authors found that people are slower in identifying happy and angry faces than they are in identifying faces with neutral expression.

The proposed framework involves an initial assessment of the expression of an unknown face, and uses that assessment to facilitate its recognition. The incoming 3D range image is processed by an expression recognition system to find the most appropriate expression label for it. The expression labels include the six prototypical expressions of the faces, which are happiness, sadness, anger, fear, surprise and disgust[5], plus the neutral expression. Therefore the expression recognition system will assign one of the seven expressions to the incoming

face. According to different expressions, a matching face recognition system is then applied. If the expression is recognized as neutral, then the incoming 3D range image is directly passed to the neutral expression face recognition system, which uses the features of the probe image to directly match those of the gallery images, which are all neutral, to get the closest match. If the expression found is not neutral, then for each of the six prototypical expressions, a separate face recognition subsystem should be used. In these cases, the system will find the matching face through modeling the variations of the face features between the neutral face and the face with expression. Since the recognition through modeling is a more complex process than the direct matching for the neutral face, this framework would support the experimental findings in [4]. Figure 1 shows a simplified version of this framework. This simplified diagram only deals with the smiling expression, which is the most commonly displayed by people publicly.

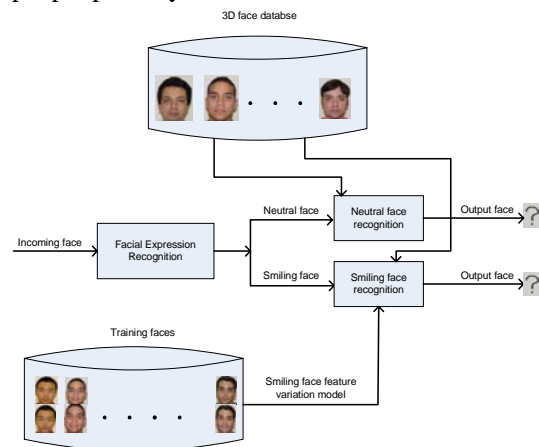


Figure1 Simplified framework of 3D face recognition

3 Data Acquisition and Preprocessing

To test the idea proposed in this model, a database, which includes 30 subjects, was built. In this database, we test the different processing of the two most common expressions, i.e., smiling versus neutral. Each subject participated in two sessions of the data acquisition process, which took place in two different days. In each session, two 3D scans were acquired. One was a neutral expression; the other was a happy (smiling) expression. The 3D scanner used was a Fastscan 3D scanner from Polhemus Inc. [6]. The resulting database contains 60 3D neutral scans and 60 3D smiling scans of 30 subjects. The left image in Figure 2 shows an example of the 3D scans obtained using this scanner, the right image is the

2.5D range image used in the algorithm, which was obtained by preprocessing as described in [8].

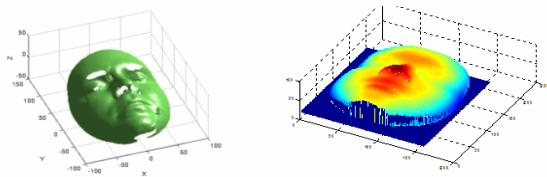


Figure 2 3D surface (left) and mesh plot of the converted range image (right)

4 Expression Recognition

Facial expressions constitute a basic mode of nonverbal communication among people. In [5], Ekman and Friesen proposed six primary emotions. Each possesses a distinctive content together with a unique facial expression. They seem to be universal across human ethnicities and cultures. These six emotions are happiness, sadness, fear, disgust, surprise and anger. Together with the neutral expression, they also form the seven basic prototypical facial expressions.

Automatic facial expression recognition has gained more and more attention recently. It has various potential applications in improved intelligence in human computer interfaces, image compression and synthetic face animation. As in face recognition, most contemporary facial expression systems use two-dimensional images or videos as data format. Logically, the same 2D shortcomings will hamper 2D expression recognition, (i.e., 2D formats are dependent on the pose of the subjects and on the illumination of the environment).

In our experiment, we aim to recognize social smiles, which were posed by each subject. Smiling is generated by contraction of the zygomatic major muscle. This muscle lifts the corner of the mouth obliquely upwards and laterally, producing a characteristic “smiling expression”. So, the most distinctive features associated with the smile are the bulging of the cheek muscle and the uplift of the corner of the mouth, as shown in Figure 3.

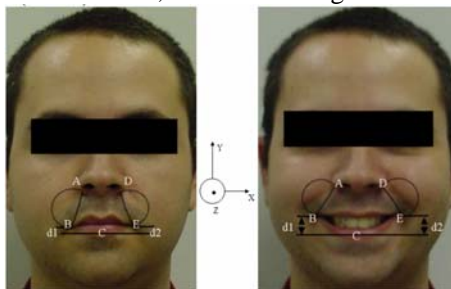


Figure 3 Illustration of features of a smiling face

The following steps are followed to extract six representative features for the smiling expression:

1. An algorithm is developed to obtain the coordinates of five characteristic points A, B, C, D and E in the face range image as shown in Figure 3. A and D are at the extreme points of the base of the nose. B and E are the points defined by the corners of the mouth. C is in the middle of the lower lip.
2. The first feature is the width of the mouth, BE, normalized by the length of AD. Obviously, while smiling the mouth becomes wider. The first feature is represented by mw .
3. The second feature is the depth of the mouth (The difference between the Z coordinates of points BC and EC) normalized by the height of the nose to capture the fact that the smiling expression pulls back the mouth. This second feature is represented by md .
4. The third feature is the uplift of the corner of the mouth, compared with the middle of the lower lip $d1$ and $d2$, as shown in the figure, normalized by the difference of the Y coordinates of points AB and DE, respectively and represented by lc .
5. The fourth feature is the angle of AB and DE with the central vertical profile, represented by ag .
6. The last two features are extracted from the semicircular areas shown, which are defined by using AB and DE as diameters. The histograms of the range (Z coordinates) of all the points within these two semicircles are calculated.

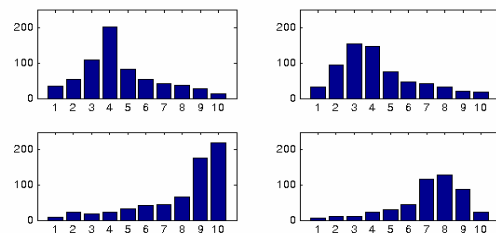


Figure 4 Histogram of range of cheeks (L &R) for neutral (top row), and smiling (bottom row) face

Figure 4 shows the histograms for the smiling and the neutral faces of the subject in Figure 3. The two figures in the first row are the histograms of the range values for the left cheek and right cheek of the neutral face image; the two figures in the second row are the histograms of the range values for the left cheek and right cheek of the smiling face image.

From the above figures, we can see that the range histograms of the neutral and smiling expressions are

different. The smiling face tends to have large values at the high end of the histogram because of the bulge of the cheek muscle. On the other hand, a neutral face has large values at the low end of the histogram distribution. Therefore two features can be obtained from the histogram: one is called the ‘histogram ratio’, represented by hr , the other is called the ‘histogram maximum’, represented by hm .

$$hr = \frac{h6 + h7 + h8 + h9 + h10}{h1 + h2 + h3 + h4 + h5} \quad (1)$$

$$hm = i; \quad i = \arg\{\max(h(i))\} \quad (2)$$

In summary, six features, i.e. mw , md , lc , ag , hr and hm , are extracted from each face for the purpose of expression recognition.

After the features have been extracted, this becomes a general classification problem. Two pattern classification methods are applied to recognize the expression of the incoming faces. The first method used is a linear discriminant (LDA) classifier, which seeks the best set of features to separate the classes. The other method used is a support vector machine (SVM). For our work, Libsvm [7] was used to implement a suitable support vector machine.

5 3D Face Recognition

5.1 Neutral Face Recognition

In our earlier research work, we have found that the central vertical profile and the contour are both discriminant features for every person[8]. Therefore, for neutral face recognition, the same method as in [9] is used: the results of central vertical profile matching and contour matching are combined. The combination of the two classifiers improves the overall performance significantly. The final similarity score for the probe image is the product of ranks for each of the two classifiers (based on the central vertical profile and contour). The image with the smallest score in the gallery will be chosen as the matching face for the probe image.

Table 1 Expression recognition results

5.2 Smiling Face Recognition

For the recognition of smiling faces we have adopted the probabilistic subspace method proposed by B. Moghaddam et al. [10, 11]. It is an unsupervised technique for visual learning, which is based on density estimation in high dimensional spaces created through eigen decomposition. Using the

probabilistic subspace method, a multi-class classification problem can be converted into a binary classification problem.

In the experiment for smiling face recognition, because of the limited number of subjects (30), the central vertical profile and the contour are not used directly as vectors in a high dimensional subspace. Instead, they are down sampled to a dimension of 17 to be used. The dimension of difference in feature space is set to be 10, which contains approximately 97% of the total variance. The dimension of difference from feature space is 7.

The results of central vertical profile matching and contour matching are combined. Here also the combination of the two classifiers improves the performance. The final similarity score for the probe image is the product of ranks for each of the two classifiers. The image with the smallest score in the gallery will be chosen as the matching face for the probe image.

6 Experiments and Results

In order to evaluate the performance of the suggested framework, one gallery and three probe databases were created. The gallery database has 30 neutral faces, one for each subject, recorded in the first data acquisition session. Three probe sets are formed as follows:

Probe set 1: 30 neutral faces acquired in the second session.

Probe set 2: 30 smiling faces acquired in the second session.

Probe set 3: 60 faces, (probe set 1 and probe set 2).

Experiment 1: Testing the expression recognition module

The leave-one-out cross validation method is used to test the expression recognition classifier. Every time, the faces collected from 29 subjects in both data acquisition sessions are used to train the classifier and the four faces of the remaining subject collected in both sessions are used to test the classifier. Two classifiers are used. One is the linear discriminant classifier; the other is a support vector machine classifier. The results are shown in Table 1.

Table 1 Expression recognition results

Method	LDA	SVM
Expression recognition	90.8%	92.5%

Experiment 2: Testing the neutral and smiling recognition modules separately

In the first two sub-experiments, probe faces are directly fed to the neutral face recognition module. In the third sub experiment, the leave-one-out cross validation is used to verify the performance of the smiling face recognition module alone.

- 2.1 Neutral face recognition: probe set 1. (Neutral face recognition module used.)
- 2.2 Neutral face recognition: probe set 2. (Neutral face recognition module used.)
- 2.3 Smiling face recognition: probe set 2. (Smiling face recognition module used.)

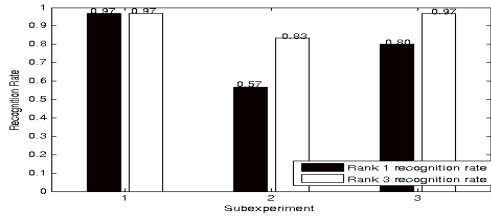


Figure 5 Results of Experiment 2(three sub-experiments)

From Figure 5, it can be seen that when the incoming faces are all neutral, the algorithm which treats all the faces as neutral achieves a very high rank-one recognition rate (97%). On the other hand, if the incoming faces are smiling, then the neutral face recognition algorithm does not perform well, only 57% rank-one recognition rate is obtained. In contrast, when the smiling face recognition algorithm is used to deal with smiling faces, the recognition rate can be as high as 80%.

Experiment 3: Testing a practical scenario

These sub-experiments emulate a realistic situation in which a mixture of neutral and smiling faces (probe set 3) must be recognized. Sub-experiment 1 investigates the performance obtained if the expression recognition front end is bypassed, and the recognition of all the probe faces is attempted with the neutral face recognition module alone. The last two sub-experiments implement the full framework shown in Figure 1. In 3.2 the expression recognition is performed with the linear discriminant classifier, while in 3.3 it is implemented through the support vector machine approach.

- 3.1 Neutral face recognition module used alone: probe set 3 is used
- 3.2 Integrated expression and face recognition: probe set 3 is used. (Linear discriminant classifier for expression recognition.)
- 3.3 Integrated expression and face recognition: probe set 3 is used. (Support vector machine for expression recognition.)

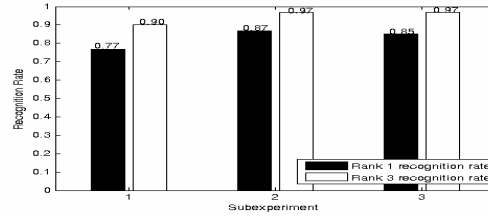


Figure 6 Results of Experiment 3 (three sub-experiments)

It can be seen in Figure 6 that if the incoming faces include both neutral faces and smiling faces, the recognition rate can be improved about 10 percent by using the integrated framework proposed here.

7 Discussion and Conclusion

Experiment 1 was aimed at determining the level of performance of the Facial Expression Recognition Module, by itself. Using the leave-one-out cross validation approach, 30 different tests were carried out (Each using 29 x 2 neutral faces and 29 x 2 smiling faces for training). The average success rate in identifying the expressions of the face belonging to the subject not used for training, in each case, was 90.8% with LDA and 92.5% when SVM was used. This confirms the capability of this module to successfully sort these two types of faces (neutral vs. smiling). Both algorithms were applied on the six facial features obtained from the range images (*mw*, *md*, *lc*, *ag*, *hr* and *hm*). Using these features, the actual choice of algorithm used to separate neutral from smiling faces did not seem to be critical.

Experiment two was carried out to test one of the basic assumptions behind the framework proposed (Figure 1). That is, a system meant to recognize neutral faces may be successful with faces that are indeed neutral, but may have much less success when dealing with faces displaying an expression, e.g., smiling faces. This differentiation was confirmed by the high rank-one recognition (97%) achieved by the Neutral Face Recognition Module for neutral faces (probe set 1) in sub-experiment 1, which was in strong contrast with the much lower rank-one recognition rate (57%) achieved by this same module for smiling faces (probe set 2), in sub-experiment 2. On the other hand, in the third sub-experiment we confirmed that a module that has been specifically developed for the identification of individuals from smiling probe images (probe set 2) is clearly more successful in this task (80% rank-one recognition).

Finally, Experiment 3 was meant to simulate a more practical scenario, in which the generation of probe images does not control the expression of the

subject. Therefore for all three sub-experiments in Experiment 3 we used the comprehensive probe set 3, including one neutral range image and one smiling range image from each of the subjects. In the first sub-experiment we observe the kind of results that could be expected when these 60 probe images are processed by a "standard" Neutral Face Recognition Module alone, which is similar to several of the contemporary approaches used for 3D face recognition. Unfortunately, with a mix of neutral and smiling faces this simple system only achieves a 77% rank-one face recognition (much lower than the 97% obtained for probe set 1, made up of just neutral faces, in Experiment 2). This result highlights the need to account for the possibility of a non-neutral expression in 3D face recognition systems. On the other hand, in sub-experiments two and three we apply the same mixed set of images (Probe set 3) through the complete process depicted in our proposed framework (Figure 1). That is, every incoming image is first sorted by the Facial Expression Recognition Module and accordingly routed to either the Neutral Face Recognition Module or the Smiling Face Recognition Module, where the identity of the subject is estimated. The right-most four columns in Figure 6 show that, whether using the linear discriminant analyzer or the support vector machine for the initial expression sorting, the rank-one face recognition levels achieved by the overall system are higher (87%, 85%).

In reviewing the results of these experiments, it should be noted that all the experiments involving smiling faces are done using the leave-one-out cross validation method because of the size of the database. Therefore the results displayed are the average, not the best one. For simplicity of implementation, the training samples for the expression recognition system and the smiling face recognition systems are the same faces. In a real application, we would select the training samples to make the best classifier for expression recognition and the identification of faces with a type of expression separately. Considerable performance improvement might be achieved in this way.

The work reported in this paper represents an attempt to acknowledge and account for the presence of expression on 3D face images, towards their improved identification. The method introduced here is computationally efficient. Furthermore, this method also yields as a secondary result the information of the expression found in the faces. Based on these findings we believe that the acknowledgement of the impact of expression on 3D face recognition and the development of systems that account for it, such as the framework introduced

here, will be keys to future enhancements in the field of 3D Automatic Face Recognition.

8 Acknowledgement

This work was sponsored by NSF grants IIS-0308155, CNS-0520811, HRD-0317692 and CNS-0426125. Mr. Chao Li is the recipient of an FIU Dissertation Year Research Fellowship.

References:

- [1] R. Chellappa, C. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proceedings of the IEEE*, 1995. **83(5)**: pp. 705-740.
- [2] W. Zhao, R. Chellappa, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Survey*, 2003. **35**: pp. 399-458.
- [3] K. Bowyer, K. Chang, and P. Flynn, "A Survey of Approaches to 3D and Multi-Modal 3D+2D Face Recognition," *IEEE International Conference on Pattern Recognition*, 2004. pp. 358-361.
- [4] N. Etcoff, J. Magee, "Categorical perception of facial expressions," *Cognition*, 1992. **44**: pp. 227-240.
- [5] P. Ekman, W. Friesen, "Constants across cultures in the face and emotion," *Journal of Personality and Social Psychology*, 1971. **17(2)**: pp. 124-129.
- [6] www.polhemus.com.
- [7] C. Chang, C. Lin, LIBSVM: a library for support vector machines, 2001. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- [8] C. Li, A. Barreto, "Profile-Based 3D Face Registration and Recognition," *Lecture Notes on Computer Science*, 2005. **3506**: pp. 484-494.
- [9] C. Li, A. Barreto, J. Zhai and C. Chin. "Exploring Face Recognition Using 3D Profiles and Contours," *IEEE SoutheastCon 2005*. Fort Lauderdale. pp. 576-579.
- [10] B. Moghaddam, A. Pentlind, "Probabilistic Visual Learning for Object Detection," *International Conference of Computer Vision (ICCV'95)*, 1995. pp. 786-793.
- [11] B. Moghaddam, A. Pentlind, "Probabilistic Visual Learning for Object Representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997. **19(7)**: pp. 696-710.