BIOMETRIC IDENTIFICATION USING 3D FACE SCANS

Chao Li   Armando Barreto   Craig Chin   Jing Zhai
Electrical and Computer Engineering Department
Florida International University
Miami, Florida, 33174, USA

ABSTRACT

Biometrics is an emerging area of bioengineering that pursues the characterization of a person by means of something that the person is or produces. Face recognition is a particularly attractive biometric challenge. Most of the face recognition research performed in the past used 2D intensity images. However, algorithms based on 2D images are not robust to changes of illumination in the environment or orientation of the subject. The ability to acquire 3D scans of human faces removes those ambiguities, since they capture the exact geometry of the subject, invariant to illumination and orientation changes. Unencumbered by those limitations, research in 3D face recognition is now beginning to address a different source of error in biometric recognition: facial geometry deformation caused by facial expressions, which can make 3D algorithms which treat 3D faces as rigid surfaces fail. In this paper, a 3D face recognition framework is proposed to tackle this problem. The framework is composed of three subsystems: expression recognition system, expressional face recognition system and neutral face recognition system. In particular, 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.
Keywords: 2D, 3D, biometrics, contour, expression, face, intensity, profile, range, recognition, registration

INTRODUCTION

Biometrics is the emerging area of bioengineering that deals with recognition of a person through something the person has (e.g., finger print, iris, retina, face, etc.) or produces (e.g., voice, handwriting, gait, etc.). The importance of biometric research has increased significantly in the new century due to its potential impact to boost security efforts. Biometric identification would also be useful in areas such as ATM machine access control, advanced human-computer interfaces, etc.

Biometric face recognition, otherwise known as Automatic Face Recognition (AFR), is a particularly attractive biometric approach [1], since it focuses on the same identifier that humans use primarily to distinguish one person from another: their faces. Much of the previous research on face recognition used 2D intensity images, and significant results have been achieved with them. However, most of the 2D face recognition systems are sensitive to the illumination changes or orientation changes of the subjects. All these problems result from the incomplete information contained in a 2D image about a face. On the other hand, a 3D scan of a subject’s face has complete geometry information about the face. It is believed that, on average, 3D face recognition methods will achieve higher recognition rates than their 2D counterparts [2]. With the rapid development of 3D imaging technology, 3D face recognition will become more and more popular, adding to the several efforts that have been undertaken to date [3]. While 3D face recognition can be robust against illumination or subject orientation variations, it is still challenged by the potential deformation of the face surface that may be present if the subject displays an emotional expression (e.g., smiles or frowns) during the 3D scan acquisition. If these potential variations are not considered, systems that treat a face as a rigid surface are prone to fail when dealing with faces with expressions. In this paper, we propose an approach to tackle the expression challenge in 3D face recognition by integrating the classification of expressions (which would default to “neutral”) in the face recognition framework.

METHODS

Our proposed framework acknowledges the potential presence of an expression in a 3D facial scan that must be identified. The framework involves an initial assessment of the expression of an unknown face. So, the process starts with an expression recognition system sorting incoming 3D range images according to the expression they contain. Incoming faces could be identified as having one of the six prototypical expressions described in [4]: happiness, sadness, anger, fear, surprise and disgust, plus the neutral expression. The framework then calls for the use of a face recognition module that is appropriate for the expression found in the incoming face. If the expression is recognized as neutral, then the incoming 3D range image is directly passed to the neutral face recognition system, which uses the features of the probe image to directly match those of the (neutral) gallery images, to get the closest match. Otherwise, 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. 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.

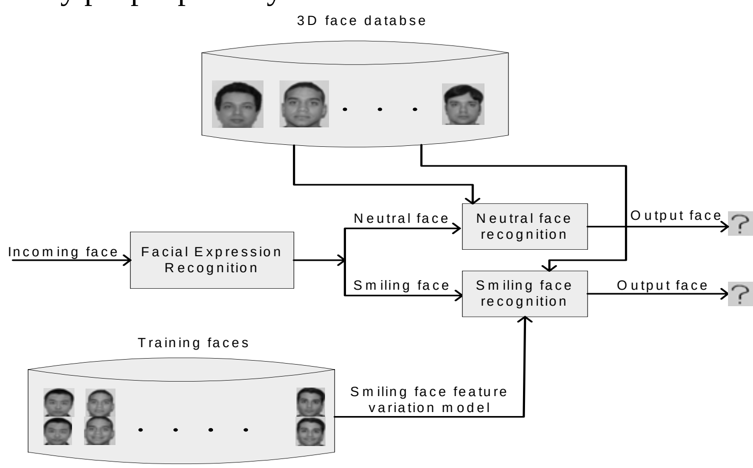

Fig 2 Simplified framework of 3D face recognition

To test the idea proposed in this model, a database, which includes 30 subjects, was built. Using 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. In each session, two 3D scans were acquired. One was a neutral expression; the other was a happy (smiling) expression. The resulting database contains 60 3D neutral scans and 60 3D smiling scans of 30 subjects. Figure 2 shows a sample range image as a mesh plot (left), and as a gray level plot (right).

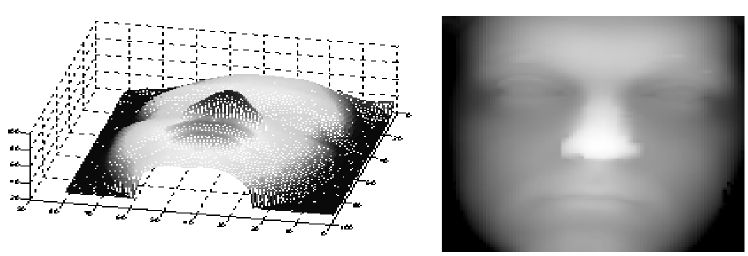

Fig 2 Mesh plot of the range data (left) and gray level image of the range data (right)

Expression Recognition

In [4], Ekman and Friesen proposed six primary emotions. Each possesses a distinctive content together with a unique facial expression. 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. The purpose of facial expression recognition is to categorize the face into one of the six expressions (or state that the face is neutral).

Specifically, in our experiment, we aim to recognize social smiles, which were posed by each subject. Smiling involves the contraction of the zygomatic major muscles, which lift the corners of a mouth obliquely upwards and laterally, producing a characteristic "smiling expression", accompanied by bulging of the cheeks (Figure 3, right panel).

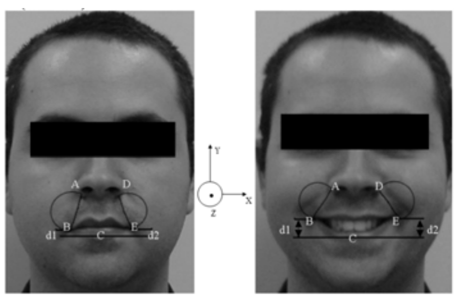

Fig 3 Illustration of features of a smiling face

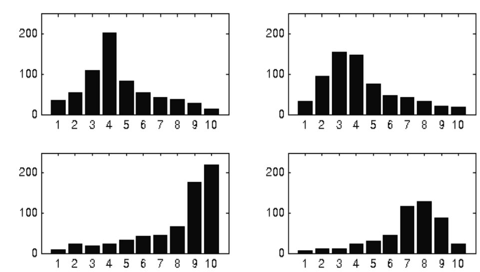

Fig 4 Histogram of range of cheeks (L &R) for neutral (top row), and smiling (bottom row) face

Our procedure to determine if a face is displaying a smiling expression entails:

- 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.
- Determine feature "*mw*", the width of the mouth, BE, normalized by the length of AD. Obviously, while smiling the mouth becomes wider.
- Determine feature "*md*", as 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.
- Determine feature "*lc*", as 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.
- Determine feature "*ag*", as the angle of AB and DE with the central vertical profile.
- Extract features "*hr*" and "*hm*" from the semicircular areas shown, which are defined by using AB and DE as diameters. The process to define these features is described below.

Figure 4 shows the histograms (Z coordinates) for the smiling and the neutral faces of the subject in Figure 3, within the semicircular areas outlined. 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 second row shows the corresponding histograms for 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*:

$$hr = \frac{h6 + h7 + h8 + h9 + h10}{h1 + h2 + h3 + h4 + h5} \qquad (1) \qquad\qquad hm = i\;; \quad i = \arg\{\max(h(i))\} \qquad (2)$$

- 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), which has been successfully used for very diverse pattern classification problems [5, 6]. For our work, Libsvm [7] was used to implement a suitable support vector machine.

3D Face Recognition

Neutral Face Recognition

In previous 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 to improve 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 gallery image with the smallest product will be chosen as the matching face for the probe image.

Smiling Face Recognition

The probabilistic subspace method proposed by B. Moghaddam et al.[10, 11] has been adapted for the recognition of smiling faces. This is an unsupervised technique for visual learning, which is based on density estimation in high dimensional spaces created by means of eigen-decomposition. Through its use, a multi-class classification problem can be converted into a binary classification problem.

Due to the limited number of subjects (30) in our experiment, the samples forming the central vertical profile and the contour are not used directly as vectors in a high dimensional subspace for smiling face recognition. 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. In this case, 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.

RESULTS

The samples used for evaluation of the framework were organized as one gallery and three probe databases. 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 together).

The validation experiments were organized as follows:

Experiment 1: Testing the neutral and smiling recognition modules separately

1.1 Neutral face recognition: probe set 1. (Neutral face recognition module used.)
1.2 Neutral face recognition: probe set 2. (Neutral face recognition module used.)
1.3 Smiling face recognition: probe set 2. (Smiling face recognition module used.)

Experiment 2: Testing a practical scenario

2.1 Neutral face recognition module used alone: probe set 3 is used
2.2 Integrated expression and face recognition: probe set 3 is used. (Linear discriminant classifier is used for expression recognition.)
2.3 Integrated expression and face recognition: probe set 3 is used. (Support vector machine is used for expression recognition.)

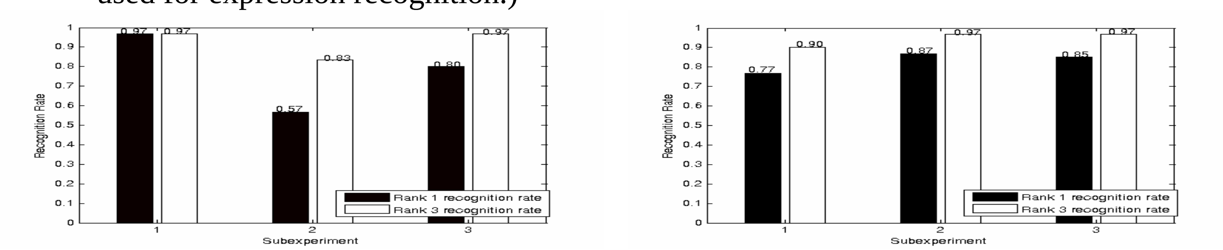

Fig 5 Results of Experiment 1 (left) & 2 (right)

DISCUSSION

Experiment 1 tested one of the basic assumptions behind the framework proposed (Figure 1). It was expected that a system designed to recognize neutral faces would be successful with faces that are indeed neutral, but it may achieve much less success when dealing with faces displaying an expression, (e.g., smiling faces). These expectations were 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, and the much lower rank-one recognition rate (57%) achieved by this same module for smiling faces (probe set 2), in sub-experiment 2. In contrast, the third sub-experiment 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).

Experiment 2 simulated a more realistic scenario, in which the expression in the subject is not controlled. Accordingly, for all three sub-experiments in Experiment 2 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. It was observed that with a mix of neutral and smiling faces this simple system only achieves 77% rank-one face recognition. 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). The right-most four columns in Figure 5 show that, whether using the linear discriminant analysis classifier 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%).

CONCLUSIONS

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. It is evident that facial expressions can produce significant deformations in the face surface. As such, facial expressions may cause deterioration in the performance of 3D face recognition systems that have been developed to process neutral faces. This was, in fact, the case when a classifier attempted to match smiling probe faces to neutral faces in a gallery (experiment 1.2). This paper presented a framework in which an incoming probe face would first be processed to detect which kind of expression, if any, might be present in it, so that this extra piece of information can be used to employ an adequate specialized face recognition sub-system for its identification. As shown in this paper, we have tested the feasibility of the framework proposed by developing a minimal implementation of it, in which the only non-neutral expression considered is for smiling faces. Notwithstanding the simplicity of this implementation, it served the purpose of demonstrating that the rank-one recognition of a mixture of neutral and smiling faces was about 10% higher when the complete structure of the proposed framework was used, i.e., when an expression recognition sub-system sorted the incoming faces and channeled them to matched recognition modules (i.e., one for neutral faces and a different one for smiling faces).

ACKNOWLEDGMENTS

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] A.K.Jian, R.Bolle, S.Pankanti(Eds), “Biometrics-personal identification in networked society” 1999, Norwell, MA: Kluwer.

[2] C.Hesher, A.Srivastava, G.Erlebacher, “A novel technique for face recognition using range images” in the Proceedings of Seventh International Symposium on Signal Processing and Its Application, 2003.

[3] K. Bowyer, K.Chang, P. Flynn, “A survey of approaches to 3D and multi-modal 3D+2D face recognition” in IEEE International Conference on Pattern Recognition, 2004: pages 358-361.

[4] P.Ekman, W. Friesen, “Constants across cultures in the face and emotion,”in Jounal of Personality and Social Psychology, 1971. **17**(2): pages 124-129.

[5] M. Adjouadi and A. Zong, "Multidimensional Pattern Recognition and Classification of White Blood Cells Using Support Vector Machines", Journal of Particle and Particle Systems Characterization, Wiley-VCH, Volume 22, Issue 2**,** pp. 107-118, September 2005.

[6] M. Adjouadi and M. Ayala, “Introducing Neural Studio: An Artificial Neural Networks Simulator for Educational Purposes”, Computers in Education Journal, Vol. 14, No. 3, pp. 33-40, July-Sept. 2004..

[7] C. Chang, C. Lin, LIBSVM: a library for support vector machines, 2001.

[8] C.Li, A.Barreto, “Profile-Based 3D Face Registration and Recognition”. in Lecture Notes on Computer Science, 2005. 3506: pages 484-494.

[9] C.Li, A.Barreto, J.Zhai, C.Chin, “ Exploring Face Recognition Using 3D Profiles and Contours” in the Proceedings of IEEE SoutheastCon 2005: pages 576-579.

[10] B. Moghaddam, A. Pentlend, “Probabilistic Visual Learning for Object Representation,” in IEEE Trans. on Pattern Analysis and Machine Intelligence, 1997. 19(7): pages 696-710.

[11] B. Moghaddam, A. Pentlend, “Probabilistic Visual Learning for Object Detection,” in the Proceedings of International Conferrence of Computer Vision (ICCV' 95), 1995: pages 786-793.