Exploring Face Recognition by Combining 3D Profiles and Contours

Chao Li Armando Barreto Jing Zhai Craig Chin
Electrical and Computer Engineering Department
Florida International University
Miami, Florida, 33174, USA
{cli006, barretoa, jzhai002,cchin006}@fiu.edu

Abstract

Most of the face recognition research performed in the past used 2D intensity images obtained by a photographic camera as the data format for processing. But the algorithms developed based on 2D images are not robust to changes of the conditions in which the images are taken, like the illumination of the environment and the orientation of the subject. With the development of 3D imaging techniques, 3D face recognition is becoming a natural choice to overcome the shortcomings of 2D face recognition, since a 3D face image records the exact geometry of the subject, invariant to illumination and the orientation changes. In this paper, a new algorithm for automatic face recognition, based on the characterization of faces by their contours and profiles, is proposed. Experiments show that the central vertical profile and the contour are both very useful features for face recognition. When combined, better recognition rates can be obtained than just using any of them alone. The performance of the algorithm is also compared with that of the traditional principal component analysis method using a database of 80 subjects. Results show that our method, which characterizes a face through its central vertical profile and contour, can achieve better results and requires less computational power in processing this test database.

1. Introduction

Automatic face recognition has been widely studied during the last few decades. It is an active research area spanning many disciplines such as image processing, pattern recognition, computer vision, neural networks, artificial intelligence, biometrics. Many researchers from these different disciplines work toward the goal of endowing machines or computers with the ability to recognize human faces as we human beings do, effortlessly, in our everyday life.

Face recognition has a wide range of potential applications for commercial, security, and forensic purposes. These applications include automated crowd surveillance, access control, mug shot identification (e.g., for issuing driver licenses), credit card authorization, ATM machine access control, design of human computer interfaces, etc.

Compared with other biometric characteristics, like finger print, iris, palm print, etc., the face is considered to be the most immediate and transparent biometric modality. Despite its intrinsic complexity, face-based recognition still remains of particular interest because it is perceived psychologically and physically as noninvasive [9].

The definition of face recognition was formulated in [6] as: "Given an image of a scene, identify one or more persons in the scene using a stored database of faces." 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.
- **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.

Until recently, most of the research work done in the area of face recognition used two-dimensional images, i.e., gray level images taken by a camera. Many new techniques emerged in this field and achieved good recognition rates. 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 geometric information about the face. It is believed that, on average, 3D face recognition methods will achieve higher recognition rates than their 2D counterparts[8]. With the rapid development of 3D imaging technology, 3D face recognition will attract more and more attention.

In [1], Bowyer provides a survey of 3D face recognition technology. Some of the techniques are derived from 2D face recognition, such as the Principal Component Analysis (PCA) to extract features from faces. Some of the techniques are unique to 3D face recognition, such as the geometric matching method, the profile matching method and the isometric transformation method [4].

This paper outlines a new algorithm used to register 3D face images automatically. Specific profiles and contours are defined in the registered faces and these are used for matching against similar features from the faces on a database including 80 subjects. The performance of the new matching algorithm is compared with that of the principal component analysis method.

2. 3D Face Database

Unlike 2D face recognition research, for which there are numerous public databases available, there are only a few 3D face databases available to researchers. Examples are the Biometrics Database from the University of Notre Dame[2] and the University of South Florida face database[3]. In our experiment, range images from the University of Notre Dame database were used.

The University of Notre Dame 3D face database includes a total of 275 subjects. 200 subjects participated in both a gallery acquisition and a probe acquisition. The time lapse between the acquisitions of the probe image and the gallery image for any subject ranges between one to thirteen weeks. In our experiment, a subset of randomly selected 80 subjects is used. For each subject, the range image which was acquired earlier is used as the gallery image; the one acquired later is used as the probe image.

The 3D scans in the UND database were acquired using a Minolta Vivid 900 range scanner. Subjects stood approximately at 1.5 meter from the camera. All subjects were asked to display a neutral facial expression and to look directly at the camera. The Minolta Vivid 900 uses a projected light stripe to acquire triangulation-based range data. The result is a 640 by 480 array of range data[5].

3. Preprocessing

In 3D face recognition, registration is a key pre-processing step. Gordon used Principal Curvature and Gaussian Curvature to segment the face surface and register it [7]. The disadvantage of using curvatures to register faces is that this process is very computationally intensive and requires very accurate range data to be successful. On the other hand, the range images in the database used for the research described in this paper contain noise, such as holes and spikes, which are difficult to eliminate fully by automatic preprocessing algorithms.

Another registration method often used involves choosing several user-selected landmark locations on the face, such as the tip of the nose, the inner and outer corners of the eyes, and then using the affine transformation to register the face to a standard position [11]. However, this approach clearly relies on human intervention, i.e., is not automatic.

A third method performs registration by using moments. The matrix constituted by the six second moments of the face surface (eq. 1) contains the rotational information of the face. By applying the Singular Value Decomposition (eq. 2), the resulting unitary matrix U can be used as an affine transformation matrix on the original face surface. The problem with this method is that during repeated scans, besides the changes in the face area for the same subject, there are also some changes outside the face area. These additional changes will also impact the registration of the face surface, causing the registration for different instances of the same subject to be different.

$$\mathrm{M} = \begin{bmatrix} m200 & m110 & m101 \\ m101 & m020 & m011 \\ m101 & m011 & m002 \end{bmatrix} \qquad (1)$$

$$U\Delta \mathrm{U}^{\mathrm{T}} = SVD(M) \qquad (2)$$

In our registration process, we assume that each subject kept his head upright during scanning, so that the face orientation around the Z axis (yaw) does not need to be corrected, but the orientation changes in the X (roll) and Y (pitch) axes need to be compensated for.

Before registration, original range images are filtered to get rid of most of the holes and the spikes. The registration process starts by locating the tip of the nose. Then a unitary transformation matrix is used to rotate the range image a certain angle around the Y (pitch) axis, and a score which measures the difference between the left and right halves of the face is computed. The angle which minimizes the score is chosen to register the face around the Y axis (pitch). This criterion is based on the symmetric property of the face.

The registration process around the X axis (roll) is similar, the transformation matrix is used to rotate the range image a certain angle around the X (roll) axis and a score is computed for each angle. But in this case the score measures the difference between two feature points in the central vertical profile. The aim is to make these two feature points, one in the forehead, and the other at the end of the chin in the central vertical profile, be at the same height.

After registration, because of the potential for different scales in the range images for the same subject, the range images are normalized, resulting in the same length for the nose bases for all the subjects. During the registration and normalization process, the bilinear interpolation method is used to find the range of a point located between two sampling points. The result of the preprocessing is a range image which has a grid of 201 by 201 points. The point (101, 101) of the grid is made to coincide with the tip of the nose in the registered face surface. The value associated with each point in the grid is the distance between the point in the face surface and the corresponding location on the XY plane. The values are offset so that the value corresponding to the tip of the nose is normalized to 100 mm. Values below 20 mm in the grid area are thresholded to 20 mm. Figure 1 shows a mesh plot of the range image and the gray level image of the same subject.

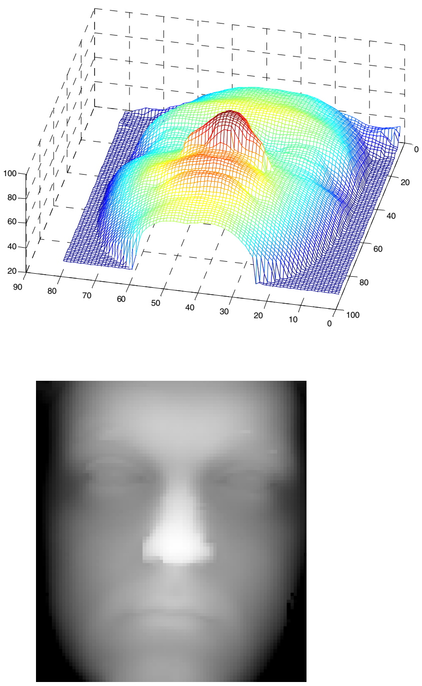

Fig 1 Mesh plot of the range image (top) and gray level image plot of range data (bottom)

4. Recognition Experiments and Results

For the experiments described here, a gallery database of 80 range images of 80 subjects (one for each subject) and a probe database of 80 different scans of the same 80 subject were used.

The use of profile matching as a means for face recognition is a very intuitive idea that has been proposed in the past [11]. In our previous research, we have, in fact, tested the efficiency of several potential profile combinations used for identification. The central vertical profile was found to be the most powerful one to characterize a face[10]. Besides profiles, the contour was also tested for its potential applicability for face recognition. In our experiment, a contour, which is defined 30mm below the tip of the nose, was extracted for each scan. Although in computing the distance or dissimilarity between profiles and contours, some researchers have used the Hausdoff distance, we found that the Euclidean distance is suitable for the context of our experiment.

The following four different approaches were tried to compare faces in the experimental data described above:

a) Central vertical profile.
b) Contour, defined 30mm below the tip of the nose.
c) Central vertical profile and contour combined.
d) Principal component analysis method.

In (a) and (b), the central vertical profile alone and the contour alone were used as the feature to compare for face recognition. In (c), the central vertical profile and the contour were combined to recognize faces. A sum rule and a multiplication rule were used in the combination of the two features. The final score for a gallery image was the sum or product of the (a) and (b) ranks, and the gallery image which had the smallest score was picked as the target subject for the probe image. The principal component analysis method was used in (d) for comparison purposes. 80 gallery range images were used for the training set. The first fifteen eigen vectors were kept to reduce the dimension of the original range images to 15. The Mahalanobis distance metric was used to find the face in the gallery.

The first approach, (a) gave a result of 70% rank-one recognition rate and (b) gave a result of 76.25% rank-one recognition rate. The rank-one recognition rates of (c) were 80% (sum rule) and 81.25% (multiplication rule), which are significantly higher than (a) and (b). The principal component analysis method in (d) gave the lowest rank-one recognition rate of only 50%.

The 3D images used for these experiments contained imperfections (e.g., holes and spikes). In spite of the application of some automatic pre-processing to eliminate some of these artifacts, a number of these imperfections were still present in the files used for the tests. For this reason, the principal component analysis method, which is very sensitive to noise, only attains a recognition rate of 50%. On the other hand, the central vertical profile method and the contour method get high recognition rates using the same data. Since these two features are extracted from part of the entire range image, they are more robust to the influence of noise. In addition, these two features are still powerful distinctive features for a face. When they are combined, the recognition rate is significantly higher than the recognition rate for any of them alone. Because these two features are poorly correlated to each other, more information benefits the recognition rate. Unlike the principal component analysis method, this method does not need training and is very simple to implement and computationally efficient.

5. Conclusion

In this paper, a new algorithm using profiles and contours to recognize faces is proposed and compared with the principal component analysis method. The following conclusions can be reached:

- Both the central vertical profile and the contour of a face have the potential to be features used in face recognition algorithms.
- Combining the central vertical profile and the contour in the algorithm yields a better recognition rate than

those using the central vertical profile or the contour alone.

- In comparison to the principal component analysis, which is sensitive to noise in the range images, methods based on the central vertical profile and the contour are more robust to the range image quality and the recognition method based on these features is simpler and less computationally intensive.

6. Acknowledgement

This work was sponsored by NSF grants IIS-0308155, HRD-0317692 and CNS-0426125. The participation of Mr. Chao Li in this research was made possible through the support of his Florida International University Presidential Fellowship.

7. References

[1] K. Bowyer, K. Chang, and P. Flynn, "*A Survey of Approaches to 3D and Multi-Modal 3D+2D Face Recognition*", in the *Proceedings of IEEE International Conference on Pattern Recognition*, 2004, pp. 358-361

[2] K. Bowyer, *University of Notre Dame Biometrics Database*, http://ww.nd.edu/%7Ecvrl/UNDBiometricsDatabase.html.

[3] K. Bowyer, S. Sarkar, *USF 3D Face Database*. 2001, http://marathon.csee.usf.edu/HumanID.

[4] A. Bronstein, M. Bronstein and R. Kimmel, "*Expression-invariant 3D face recognition"*, in the *Proceedings of Audio & Video-based Biometric Person Authentication (AVBPA)*, 2003, pp. 62-69.

[5] K. Chang, K. Bowyer and P. Flynn, "*Face recognition using 2D and 3D facial data*", in the *Proceedings of 2003 ACM Multimodal User Authentication Workshop,* 2003, pp. 25-32.

[6] R. Chellappa, C. Wilson, and S. Sirohey, "*Human and Machine Recognition of Faces: A Survey"*, in the *Proceedings of the IEEE*, 1995. **83(5)**, pp. 705-740.

[7] G. Gordon, "*Face recognition based on depth maps and surface curvature"*, in *Geometric Methods in Computer Vision, SPIE*. July 1991, pp. 1-12.

[8] C. Hesher, A.Srivastava and 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.

[9] A.K. Jian, R. Bolle and S. Pankanti(Eds.), "*Biometrics-Personal Identification in Networked Society"*, 1999, Norwell, MA: Kluwer.

[10] C. Li, A. Barreto, "*Profile-Based 3D Face Registration and Recognition*", in the *Proceedings of 7th International Conference on Information Security and Cryptology (ICISC) 2004*, pp. 456-467.

[11] T. Nagamine, T. Uemura and I. Masuda, "*3D facial image analysis for human identification*", in the *Proceedings of International Conference on Pattern Recognition (ICPR,1992)*. 1992, pp. 324-327.