A Profilometric Approach to 3D Face Reconstruction and Its Application to Face Recognition

Surath Raj Mitra and K.R. Ramakrishnan

Indian Institute of Science

Abstract. 3D Face Recognition is an active area of research for past several years. For a 3D face recognition system one would like to have an accurate as well as low cost setup for constructing 3D face model. In this paper, we use Profilometry approach to obtain a 3D face model. This method gives a low cost solution to the problem of acquiring 3D data and the 3D face models generated by this method are sufficiently accurate. We also develop an algorithm that can use the 3D face model generated by the above method for the recognition purpose.

1 Introduction

Face Recognition is a very old research problem. For the last 30 years numerous methods have been proposed by the researchers from several different fields to solve this problem [1]. Though a considerable amount of success has been achieved, still there exist a lot of unsolved questions.

A large number of techniques exist for face recognition from still images. In [2], Turk and Pentland have used that Principal Component Analysis for face recognition. Elastic bunch graph matching [3] is another interesting approach. Here each face is represented as a graph and face recognition is based on the graph similarity. In [4], Cootes et al. proposed Active appearance model (AAM) for face recognition. An AAM combines the statistical model of the appearance and shape variation in a shape normalized from.

Recently 3D face recognition has gained a significant attention among the researchers in this field. Intutively a 3D face recognition system should perform better than a 2D face recognition system as face is inherently a 3D object. In [5], Lee and Milios have used Extended Gaussian Image (EGI) based approach for 3D face recognition. Medioni and Waupotitsch [6] proposed to solve the 3D face recognition problem using iterative closest point (ICP) matching of face surfaces. A review of the 3D face recognition systems can be found in [7]. An important problem in a 3D face recognition system is the generation of an accurate 3D model. Generally laser scanners (which are very costly) are used to capture 3D model. On the other hand one can use fringe projection techniques to capture the 3D information.

Fringe projection techniques has been extensively used in industry for the purpose of 3D measurements. One of the popular fringe projection techniques

P. Kalra and S. Peleg (Eds.): ICVGIP 2006, LNCS 4338, pp. 631–640, 2006.

is Fourier Transform Profilometry(FTP) [8]. In FTP, a sinusoidal or rectangular fringe pattern is projected on to the object from an offset angle and the image of the deformed fringe pattern which is phase modulated by the topographical variations of the object surface is captured. Then the phase is extracted from the deformed fringe pattern using Fourier transform analysis. The recovered phase is limited in the interval $[-\pi, \pi]$. This is called the wrapped phase map. Phase unwrapping technique is used to get the natural phase distribution from the wrapped phase. Unwrapping process refers to adding an appropriate integer multiple of 2π to each pixel element of the wrapped phase map. Phase unwrapping should be path independent for a perfect phase map. But, local shadows or low fringe modulation, irregular surface brightness, fringe discontinuities makes phase unwrapping a path dependent problem. One popular approach for phase unwrapping is based on reliability-guided parameter map.A recent review on reliability guided phase unwrapping algorithms can be found in [9]. The unwrapped phase-map so obtained by the phase unwrapping algorithm is proportional to the height variations of the object surface.

In this paper Fourier Transform Profilometry(FTP) [8] method is used to generate a dense depth map of human face to pixel level accuracy as the phase is estimated for each pixel by a Fourier Transform. Then 2D face image estimate (i.e. the image of face without fringe pattern) is generated from the input fringe image. The 3D model of the human face along with the 2D face image estimate is then used for the purpose of the face recognition.

The rest of the paper is organized as follows: In section 2 we describe the basic theory of the Fourier Transform Profilometry(FTP). In this section we also describe the results obtained by the FTP method. In section 3 we describe the Face Recognition algorithm. In section 4 summary and the future work is discussed.

2 3D Face Reconstruction Using Fourier Transform Profilometry

In Fourier Transform Profilometry(FTP) [8],[10],[11] a sinusoidal fringe pattern is projected on a 3D object and resulting image is captured by a CCD camera. The height distribution of the three dimensional object results in phase modulation.The phase contains the information about the object profile. In FTP,Fourier analysis is used to get the phase map from the fringe image. In our experiments, we have used the human face as the three dimensional object. Though the FTP is widely used in industry, it has never been used in human 3D face reconstruction.In the next subsection we will describe the basic theory of FTP.

2.1 Theory of Fourier Transform Profilometry [8]

Takeda et al. introduced FTP in [8]. General optical geometry for Fourier transform profilometry is shown in the figure 1. This is called crossed optical axes geometry. In the figure 1, E'_pE_p is the optical axis of the projector lens and

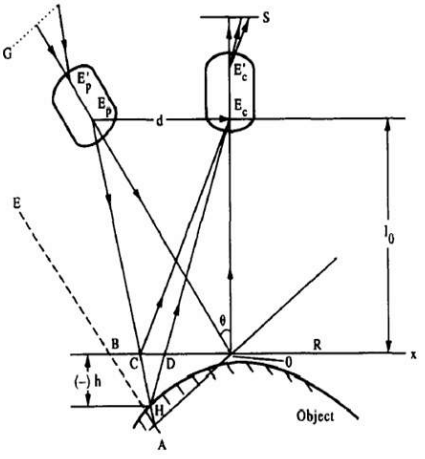


Fig. 1. Optical geometry of FTP[from [8]]

E'_cE_c is the optical axis of the camera lens and R is the reference plane. The reference plane R is a fictitious plane and serves as reference from which object height is measured. The image of a perfectly flat and uniform reference plane (i.e. the height distribution is zero) with fringe pattern projected on it can be expressed as,

$$g_0(x,y) = \sum_{n=-\infty}^{\infty} A_n \exp\{i[2\pi n f_0 x + n\phi_0(x,y)]\} \tag{1}$$

When the measured object (in our case, human face) is put on the reference plane, the deformed fringe pattern observed can be expressed as:

$$g(x,y) = r(x,y) \sum_{n=-\infty}^{\infty} A_n \exp\{i[2\pi n f_o x + n\phi(x,y)]\} \tag{2}$$

where f_0 is the fundamental frequency of the observed fringe image, $\phi(x,y)$ and $\phi_0(x,y)$ are the the phase modulation due to object height distribution and the phase modulation for reference plane (when object height distribution is zero) respectively, r(x, y) and A_n are the non uniform distribution of reflectivity on the object surface and the weighting factors of the Fourier series respectively. $\phi(x,y)$ contains the information about object profile.

In the FTP method, 1D Fourier transform of (2) is calculated. The Fourier spectra obtained is shown in the figure 2. The Fourier transform is calculated with respect to the variable x and the other variable y is treated as a fixed parameter. $r(x,y)$ and $\phi(x,y)$ are assumed to vary very slowly compared to the frequency f_0 of the fringe pattern. Using a suitable filter function only the

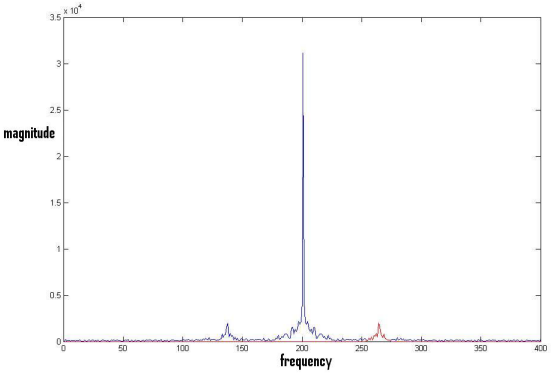


Fig. 2. Spatial frequency spectra of a deformed fringe image for a fixed value of y

fundamental spectrum is selected(shown as red in the figure 2). Applying inverse Fourier transform on the fundamental spectrum,we get,

$$g'(x,y) = A_1 r(x,y) \exp\{i[2\pi f_0 x + \phi(x,y)]\} \tag{3}$$

The same operation on (1) gives,

$$g'_0(x,y) = A_1 \exp\{i[2\pi f_0 x + \phi_0(x,y)]\} \tag{4}$$

Then the phase change ($\triangle\phi(x,y)$) is given by,

$$\triangle\phi(x,y) = \phi(x,y) - \phi_0(x,y) \tag{5}$$

From (3) and (4) $\triangle\phi(x,y)$ is obtained as,

$$\triangle\phi(x,y) = Im(\log(g'(x,y)g'^*_0(x,y))) \tag{6}$$

where * denote the conjugate operation.

$\triangle\phi(x,y)$ is approximately proportional to the height variation and thus it is the measure of the object height distribution. The phase calculated here is the wrapped phase. Using phase unwrapping algorithm, we can get the natural phase distribution. In our work we have used zpm [12] algorithm for phase unwrapping.

2.2 Results of FTP

The experimental setup consists of a LCD projector and a CCD camera. Sinusoidal fringe patterns are projected on the face of a subject using the LCD projector and the CCD camera is used to capture the image of the fringe deformation on the face. Figure 3 shows input fringe pattern which is a sinusoidal

fringe pattern. The fringe pattern we are using varies only in the vertical direction. So Fourier transform is applied on each column (scan line) of the image. Figure 4 shows different steps in FTP. Figure 4(a) shows the input fringe image. Figure 4(b) shows the wrapped phase map. The gray value at a pixel in the figure 4(b) is proportional to the phase at that pixel. Figure 4(c) shows the unwrapped phase map. Figure 4(d) shows the 3D mesh plot of $\triangle\phi(x, y)$. Figure 5 shows two views of the generated 3D model.

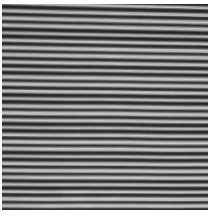
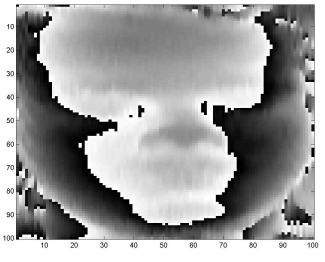
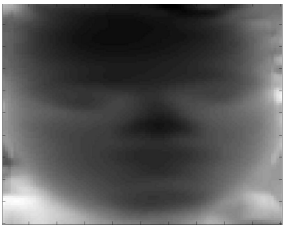


Fig. 3. Input fringe pattern

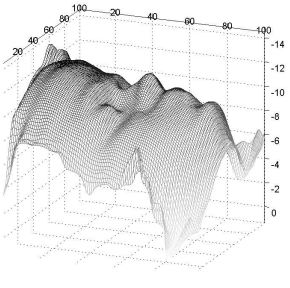
(a) input fringe image



(b) Wrapped phase map



(c) Unwrapped phase map



(d) 3D mesh

Fig. 4. Results of different steps in FTP

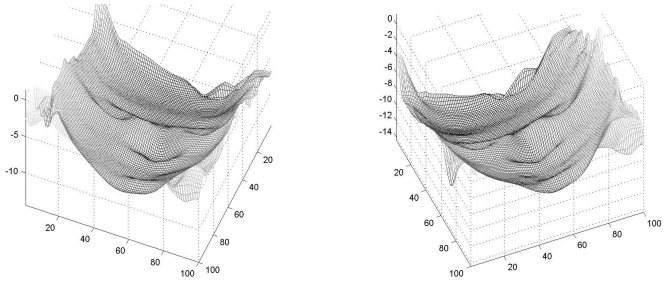


Fig. 5. Two views of generated 3D model

3 A Multimodal(2D+3D) Face Recognition System

The face recognition method we have developed uses both 2D and 3D features. The 3D feature is obtained from the 3D face model that we have generated using FTP. In order to get the 2D feature, we need to generate the 2D face image estimate(i.e. the image of face without fringe pattern) from the input face image with fringe projection. The advantage of generating the 2D image estimate is that the position of the face in the estimated image will be same as that in the 3D model. This will help us in placing the facial grid on both the 3D model and the 2D image estimate which is described latter.

3.1 Estimation of 2D Face Image

For 2D face image estimation we assume that the texture of the face does not have any repetitive patterns comparable to the frequency of the sinusoidal fringe, which in general is true. With this assumption the 2D face image estimate can be done by simple low pass filtering operation. Therefore, we need to separate the low frequency components from the high frequency components those are present in the image after fringe projection. In other words, we need to extract the envelope of the modulated sine wave. This will give us the estimate of the 2D face image. This process is done along each color channel to get a full color estimate of 2D face. Then this is converted to the gray scale as 2D features are extracted from the gray scale image. Figure 6 shows the result of the 2D image estimation.

3.2 Face Recognition Algorithm

The first step in our algorithm is the detection of the nose tip location [13]. Locating nose tip is very easy from the 3D face model. Since our 3D model gives the relative height distribution of the face, the nose tip has always maximum height. So, we just have to find the point with maximum height in the 3D face model. After locating the position of the nose tip the start and the end of the

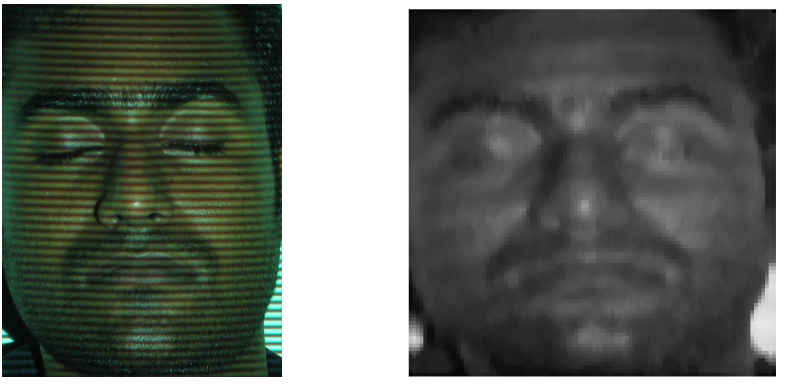


Fig. 6. Result of the 2D image estimation

nose (which approximately points to the eye location and the mouth location respectively) are found. With respect to the nose tip position a facial grid is formed to cover the eye, nose and the mouth portion of the face. Figure 7 shows a subject from our face database with facial grid on the unwrapped phase map as well as on the 2D image estimate.

In order to get the 2D feature, the 2D face image is convolved with the 2D Gabor wavelets [14][3]. The Gabor wavelets (kernels, filters) is defined as follows,

$$\Psi_{\mu,\nu}(z) = (\frac{\| k_{\mu,\nu} \|^2}{\sigma^2}) e^{-(\frac{\|k_{\mu,\nu}\|^2 \|z\|^2}{2\sigma^2})} [e^{ik_{\mu,\nu}z} - e^{-(\frac{\sigma^2}{2})}] \tag{7}$$

where μ and ν denote the orientation and scale of the Gabor kernels respectively. $z = (x, y)$, $\|\|$ denotes the norm operator, and the wave vector $k_{\mu,\nu}$ can be defined as follows,

$$k_{\mu,\nu} = k_\nu e^{i\phi_\mu}$$

where, $k_\nu = \frac{k_{max}}{f^\nu}$, $\phi_\mu = \frac{\pi\mu}{8}$ and f is the the spacing factor between kernels in the frequency domain. The estimated 2D face image is convolved with the Gabor filters.Then for all the points in the facial grid the absolute magnitudes of the complex-valued filter responses are concatenated to form a 2D feature vector (v_t).

In order to get the 3D depth/shape feature, for each point in the facial grid we have considered its eight neighborhood points. Then the mean and the variance of the local height distribution is calculated as,

$$mean = (\frac{1}{8}) \sum_{i=1}^{8} (d_n - d_i) \tag{8}$$

$$variance = (\frac{1}{8}) \sum_{i=1}^{8} (mean - (d_n - d_i))^2 \tag{9}$$

where d_i is the height of a neighborhood point and d_n is the height of the nose tip point. Here by height we mean the value of $\phi(x, y)$ at a facial grid point. So for each facial grid point we have two depth feature. The 3D feature vector (v_d) is obtained by concatenating the mean and variance feature for all the facial grid points.

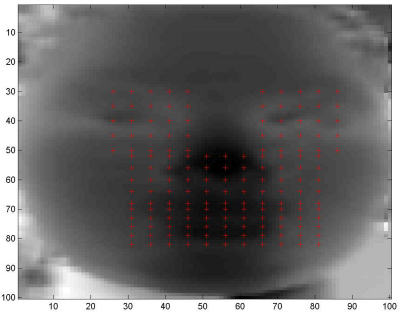
Given a test 3D face model and its 2D image estimation we form corresponding 2D feature vector (v_t) and 3D feature vector (v_d).Then the nearest neighbor classification rule is applied. In the Nearest neighbor procedure, the closest Euclidean distance match of a test face to the faces in the training set is determined. Then the level of the closest match training face is assigned to the test face. The Euclidean distance between a test face and a training face is given by,

$$D^2 = \alpha \parallel (v_t - (v_t)_k) \parallel^2 + \beta \parallel (v_d - (v_d)_k) \parallel^2 \tag{10}$$

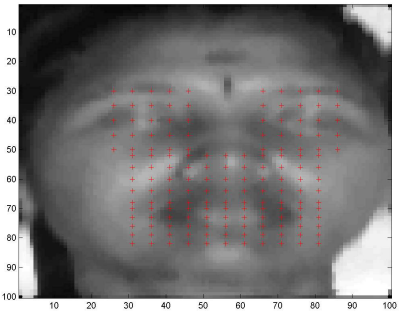
where $(v_t)_k$ and $(v_d)_k$ are the 2D and the 3D feature vector describing the k^{th} training face and α, β are suitably chosen weights.

3.3 Experiments and Results

In order to test the performance of our face recognition algorithm we have collected a database of 13 persons. Each person has two images on the database, one of which is used for training and other for testing. Each face image is collected after projecting a sinusoidal fringe pattern on the face. Fourier transform analysis is carried out on each of the captured image to get the phase distribution. This phase distribution gives us the estimate of height distribution on the face and hence the 3D model. We have also computed the 2D face image estimation for each of the captured fringe image. Then a facial grid is placed on the 2D image estimate and on the 3D model. The number of facial grid point considered is 160. Then the face recognition algorithm is applied on this database.We have used Gabor wavelets at five different scales, $\nu = \{0, ..., 4\}$and eight orientations, $\mu = \{0, ..., 7\}$. In our experiment we have used 1 and .1 as the value of α and



(a)facial grid on unwrapped phase map



(b)facial grid on 2D image estimate

Fig. 7. A subject with facial grid

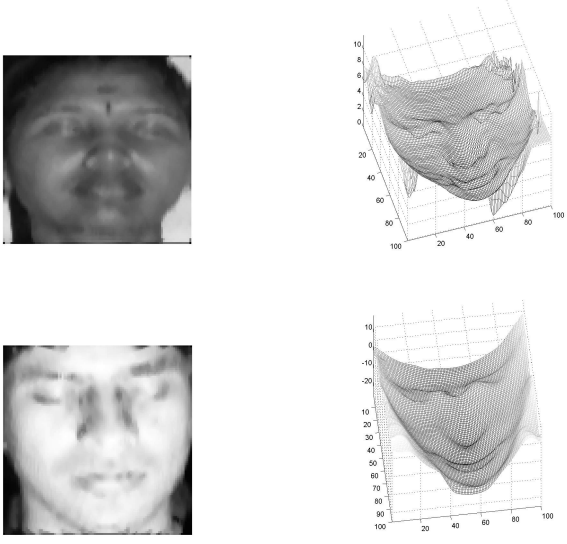


Fig. 8. 2 subjects from the face database

β respectively. An accuracy of more than 90% is obtained in our experiment. Figure 7(a) and 7(b) show a subject from our face database with facial grid on the unwrapped phase map as well as on the 2D image estimate respectively. Figure 8 shows 2 subjects from our database with corresponding 3D models.

4 Summary and Future Work

In this paper we use FTP to construct the 3D model of human face. This method needs only a LCD projector and a CCD camera to acquire the 3D data. These instruments are very cheap and easily available. Thus the FTP approach shows a way to reduce the cost associated with the 3D face data acquisition. In this paper we also propose a multimodal Face Recognition algorithm. However the algorithm is tested only on a small dataset. It needs to be tested on a larger dataset to really bring out the efficacy of this approach for face recognition. So our future work includes collection of a large amount of dataset and testing our approach on that dataset.

References

1. W. Zhao, R. Chellappa, and A. Rosenfeld, Face recognition: a literature survey, ACM Computing Surveys, vol. 35, no. 4, pp. 399-458, December 2003.
2. M. Turk and A. Pentland, Eigenfaces for recognition, Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

3. L. Wiskott, J.M. Fellous, N. Kruger, and C. von der Malsburg, Face recognition by elastic bunch graph matching, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp.775-779, 1997.
4. T. F. Cootes, G. J. Edwards and C. J. Taylor, Active appearance models, IEEE Trans. Pattern Analysis and Machine Intelligence 23, pp.681-685, 2001.
5. J. C. Lee and E. Milios, Matching range images of human faces, International Conference on ComputerVision, pp. 722-726, 1990.
6. G. Medioni and R. Waupotitsch, Face recognition and modeling in 3d. IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG 2003), p. 232.233, October 2003.
7. K. W. Bowyer, K. Chang, and P. Flynn. A Survey Of 3D and Multi-Modal 3D+2D Face Recognition, In International Conference on Pattern Recognition (ICPR) 2004.
8. M. Takeda and K. Mouth, Fourier transform profilometry for the automatic measurement of 3-d object shapes, Applied Optics, vol. 22, no. 24, pp. 3977-3982, 1983.
9. X. Su and W. Chen, Reliability guided phase unwrapping algorithm-a review, Opt and Laser Eng, vol. 42, pp. 245-261, 2004.
10. X. Su and W. Chen, Fourier transform profilometry: a review, Opt Laser Eng, vol. 35, pp. 263-284, 2001.
11. J. Li, X. Y. Su, and L. Guo, Improved fourier transform profilometry for the automatic measurement of 3d object shapes, Optical Engineering, vol. 29, no. 12, p. 1439, December 1990.
12. J. Dias and J. Leitao, The zpm algorithm for interferometric image reconstruction in sar/sas, IEEE Transactions on Image processing, vol. 11, no. 4, pp. 408-422, 2002.
13. Y. Lee, T. Yi, 3D face recognition using multiple features for local depth information, 4th EURASIP Conference focused on Video I Imaae Pmcessina and Multimedia Communications, 2-5 July 2003, Zagreb, Croatia
14. C. Liu and H. Wechsler, Independent component analysis of gabor features for face recognition, IEEE Trans. Neural Networks, vol. 14, no. 4, pp. 919-928, July 2003.