MULTIMODAL FACE RECOGNITION USING SPECTRAL COEFFICIENTS OF FACE TEXTURE IMAGES AND STATISTICAL PROCESSING OF SPECTRAL COEFFICIENTS OF FACE RANGE IMAGES

NAVEEN S
Assistant Professor, Department of ECE, LBS Institute of Technology for Women, Trivandrum, Kerala--695012, India
nsnair11176@gmail.com

DR R.S MONI
Professor, Department of ECE, Marian Engineering College, Trivandrum, Kerala-695582, India
moni2006rs@gmail.com

ABSTRACT: 3D Face recognition has been an area of interest among researchers for the past few decades especially in pattern recognition. The main advantage of 3D Face recognition is the availability of geometrical information of the face structure which is more or less unique for a subject. This paper focuses on the problems of person identification using 3D Face data. Use of unregistered 3D Face data for feature extraction significantly increases the operational speed of the system with huge database enrollment. In this work, unregistered Face data, i.e. both texture and depth is fed to a classifier in spectral representations of the same data. 2D Discrete Cosine Transform (DCT) is used here for the spectral representation. The face recognition accuracy obtained when the feature extractors are used individually is evaluated. The use of depth information alone in different spectral representation was not sufficient to increase the recognition rate. So a fusion of texture and depth statistical information of face is proposed. Application of statistical method seems to degrade the performance of the system when applied to texture data and was effective in the case of depth data. Fusion of the matching scores proves that the recognition accuracy can be improved significantly by fusion of scores of multiple representations. FRAV3D database is used for testing the algorithm.

KEYWORDS: Point Cloud, Rotation Invariance, Pose Correction, Depth Map, Spectral Transformations, CDF, Texture Map and Principal Component Analysis

INTRODUCTION

3D Face recognition has been an active area of research in the past decades. The complications encountered in the enrollment phase and the huge computational requirements in the implementation phase have been the major hindrance in this area of research. The scenario has improved tremendously due to the latest innovations in 3D imaging devices and has made 3D Face recognition system a reliable option in security systems based on Biometrics. Though poor resolution is a major drawback encountered in 3D Face images the geometrical information present in 3D facial database can be exploited to overcome the challenges in 2D face recognition systems like pose variations, bad illumination, ageing etc.

In this work, focus is made on an identification problem based on 3D Face data using fusion schemes. Identification corresponds to the person recognition without the user providing any information other than the 3D facial scan. The system arrives at an identity from among the enrolled faces in the database. Use of texture information along with the geometrical information of the face seems to improve the recognition accuracy of face recognition system when pose correction is not done as a pre-processing step. Here statistical processing of depth data is done so as to improve the recognition accuracy.

Alexander M. Bronstein et al. [2] proposed an idea of face recognition using geometric invariants using Geodesic distances. C. Beumier [3] utilized parallel planar cuts of the facial surfaces for comparison. Gang Pan et al [4] extracted ROI of facial surface by considering bilateral symmetry of facial plane. Xue Yuan et al [5] proposed a face recognition system using PCA, Fuzzy clustering and Parallel Neural networks. Trina Russ et al [6] proposed a method in which correspondence of facial points is obtained by registering a 3D Face to a scaled generic 3D reference face. Ajmal Mian et al [7] used Spherical Face Representation for identification. Ondrej Smirg et al [8] used DCT for gender classification since the DCT best describes the features after de-correlation. Hua Gao et al [9] used Active Appearance model for fitting faces with pose variations. et al [10], used 2D-PCA for getting the feature matrix vectors and used Euclidean distance for classification. , Omid Gervei et al [11] proposed an approach for 3D Face recognition based on extracting principal components of range images by utilizing modified PCA methods namely 2D-PCA and bidirectional 2D-PCA.Jain et al., [13] discusses the use of statistical methods for pattern recognition using neural networks.

A typical 3D Face is shown in Fig. 1. Fig.2 represents its axis level representation. Fig. 3 and 4 represents the texture map with two different orientations.

Figure 1: 3D Face Model

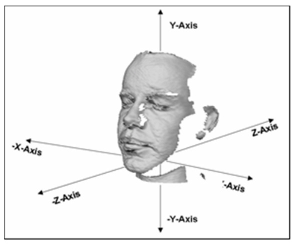


Figure 2: 3D Face in Space

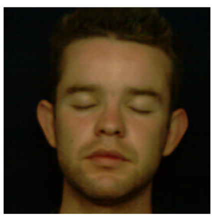


Figure 3: 2D Texture Map(Frontal View)

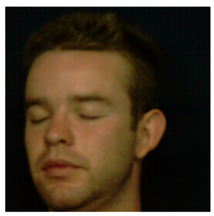


Figure 4: 2D Texture Map(Left Turned head)

Since only a sparse set of points are available in the dataset, it is necessary to increase the data density by using multiple data representations generated from same raw data. For this the data is transformed into spectral domain using DCT. This sparse set of data with occlusion can be effectively countered by invoking multiple score fusion schemes which can effectively improve the feature data density. Use of Depth information alone is not sufficient for an efficient recognition system since pose correction is not done. So texture information is also incorporated with the fusion scheme.

PROPOSED SCHEME FOR FACE RECOGNITION

The system aims at extracting the feature from the input data through feature extraction tools and fuses the scores to get a system with better recognition accuracy. The main feature extraction principle used in this system is the spectral transformation. The spectral transformation tool used here is 2D-Discrete Cosine Transform. These spectral transformations transform the data to a better representation which increases the accuracy of recognition system. Block diagram of the proposed scheme is shown in Fig. 5.

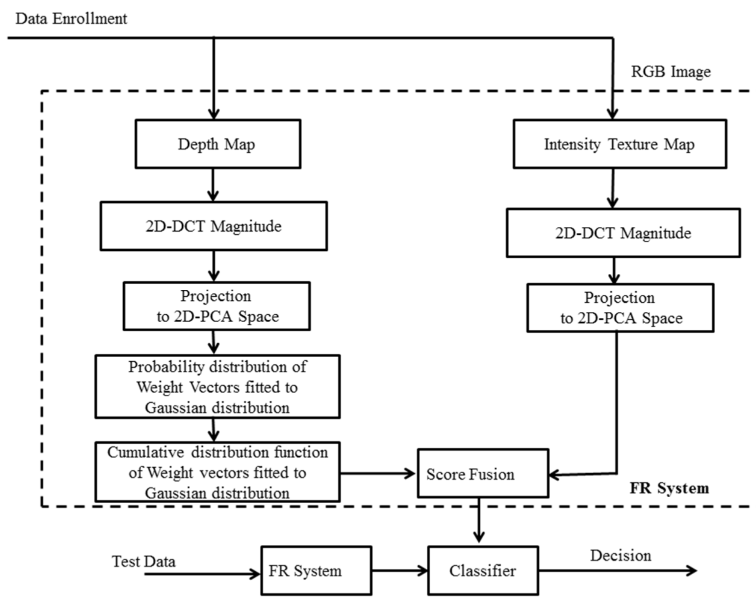


Figure 5. Proposed Method

The most important part of this work lies in the pattern classification problem. A pattern of data points is available. This pattern is not sufficient for the recognition system to work since the data will be highly occluded due to pose variations in the X, Y and Z axis or in any complex plane.

The 3D Face recognition scheme is affected by pose variations of the subject (person under consideration). There are methods available in which the correction to this effect of pose variations also is included. One such method is the Iterative Closest Point (ICP) algorithm. But the main disadvantage of these methods is that a reference face is to be used as a model for other rotated faces to be corrected. Also the processing time taken is very high. Further, the reliability of this result depends on the accuracy in selection of the reference face model used. Therefore, in this work done, this correction to the effect of pose variation is not considered. The method aims at recognizing the subject without much computationally complex mathematical procedure. Also the results prove that the efficiency of system is comparable with a system with pose correction. The idea behind spectral representation of data is that, when data is in spatial domain, comparison will be done as one to one pixel level or voxel level. So the rotation and translation of data will highly affect the result. Moreover the accuracy of the system will go down to even 5% under severe pose variations in X, Y and Z axis. When spectral transformation is done the distributed data will be concentrated or it may be represented in a more uniform way. I.e. the input data will be concentrated and represented uniformly in spectral domain. The translation and rotation invariance properties of the transformations used will aid to improve the accuracy of system significantly.

Here FRAV3D database is considered. It contains the facial data with different face orientations and expressions. When depth information alone was considered the Face recognition accuracy (FRA) was not high. So texture data of face is also considered which significantly improves the FRA.The data available for the analysis and testing will be in Depth Map format which is a matrix array of size M x N. For each face depth data input the number of depth points in 2D plane can be different.

The proposed method involves the following steps given below in sequence.

1) The 2D face depth data is first normalized with the maximum intensity value. From this 2D depth map nose tip is detected using Maximum Intensity Method and the area around the nose (ROI-Region of Interest) is extracted (Fig.6 and Fig.7).
2) On this ROI data, 2D-DCTis applied. The detailed explanations are given on following sections. Simultaneously 2D-DCT is applied over the complete face texture data.
3) Once spectral representations are obtained, Principal Component Analysis (PCA) is applied on that data to get the corresponding weight vectors.
4) Now probability distribution of the weight vectors of depth data is computed by fitting it onto Gaussian distribution. After this fitting process corresponding CDF is calculated as a new feature vector called Cumulative Depth Feature Vector (CDFV).
5) CDFV along with weight vector of texture data is fed to classifier which uses Euclidean distance for classification. Individual error scores are calculated and then these scores are fused to get the minimum score.

Depth Map Normalization

Depth map obtained is normalized with the maximum intensity value to make the depth data more visible. Here the depth values are normalized between the range 0 and 255.

$$\text{Normalized Depthmap}=\frac{\text{Original Depthmap}*255}{\text{Max } Intensity(\text{Original Depthmap })} \tag{1}$$

Nose tip Localization and face area extraction

For localizing the nose tip, maximum intensity method is used. In this method assumption is made that the nose tip will be the point with maximum pixel intensity. Once the nose tip is found the circular area (ROI) around the nose tip is extracted using an optimum radius. Now the depth map will contain the face area only, all other unwanted portions are cropped away. Next face area is centralized by making the nose tip as the center pixel of the image. Otherwise the matching process will result in a lower accuracy. The face area is also normalized by the maximum intensity. The centralized face image is as shown in Fig. 6 and Fig. 7.

Figure 6. Depth Map

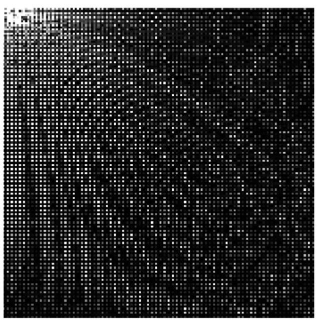
Figure 7. ROI from Depth Map

2D Discrete Cosine Transform

The depth data will be having the pixel value as the geometrical measure of the facial data. Face images have higher redundancy and pixel level correlation which is a major hindrance in face recognition systems. Transforming face images to spectral domain will reduce the redundancy. Here only the magnitude of spectral data is taken alone since it is not transformed back to spatial domain in any of the processing stages.

The depth image is transformed using 2D-DCT so that rotation effects are reduced. DCT is a rotation invariant transformation. DCT has the property of de-correlation which enables the data structure to loose spatial pixel dependency. So that the distributed pixel values (normalized) are properly aligned, this enables the pattern matching more efficient. DCT spectrum of face depth image will appear as shown in Fig. 8. Transformation to spectral domain using 2D Discrete Cosine Transformation can be done using equation (2).

$$F(U,V)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\cos\left(\frac{(2x+1)\pi u}{2M}\right)\cos\left(\frac{(2y+1)\pi v}{2N}\right), \text{ for a M x N depth image} \qquad (2)$$



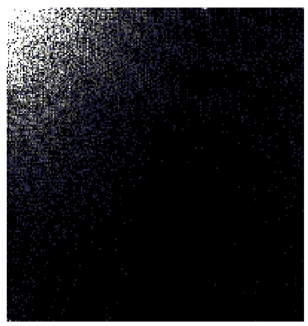


Figure 8: DCT Representation of ROI Depth Map Figure. 9: DCT Representation of Texture Map

The same procedure is repeated for the texture map also to get the spectral representation of gray scale intensity image using 2D-DCT as in Fig. 9.

Principal Component analysis

Use of spectral transformations will make the data samples almost spatially uncorrelated. Even then, some spatial dependency may exist. So Principal Component Analysis (PCA) [12] which uses the orthogonal transformations to get linear uncorrelated data sets called Principal Components is employed. Conventional covariance method is used for the above. To start with feature extraction using 1D-PCA is discussed for better understanding.

Let X_i be the spectral transformed 1D data which representations i^{th} person, it is grouped as a M' x N matrix $X=[X_1\ X_2\ \ldots X_N]$, where N is the number of face samples under consideration and M' is the length of each feature vector.

Mean vector is calculated as follows

$$X_m = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{3}$$

Standard deviation is calculated as

$$X_{SD} = \frac{1}{N}\sum_{i=1}^{N} (X_i - X_m) \tag{4}$$

Covariance matrix is calculated as

$$X_{COV} = X_{SD*} X^T_{SD} \tag{5}$$

Here, the covariance matrix is of size M' x M', which is of very large dimension. Also it gives M' Eigen values and M' Eigen vectors which are too large in number to process. Therefore, dimensional reduction is adopted by altering the construction of covariance matrix as follows.

$$X_{COV} = X^T_{SD*}\ X_{SD} \tag{6}$$

The result is a matrix of size N x N, where N is the number of subjects under consideration. It gives N Eigen values and N Eigen vectors. The Eigen values are sorted in descending order and the first N' largest Eigen values and corresponding Eigen vectors are selected as others are insignificant. Eigen vectors in N' dimension is transformed to the higher dimension of M' by multiplying with Standard deviation Matrix. The test data is projected to this lower dimension space to get the corresponding weight vectors.

Now feature extraction using 2D-PCA is considered using spectral representation of depth map. The only difference with 1D-PCA in calculating the Covariance matrix is that here a 2D matrix is used when compared to 1D Matrix in 1D-PCA. After determining the Eigen values and Eigen vectors, a 2D weight vector matrix is obtained which is then converted to a column matrix.

Cumulative Depth Feature Vector

Projection of depth data Cosine transform spectral coefficients on 2D PCA space will give weight vectors. These weight vectors are fitted on to the probability distribution function of a Gaussian distribution. The mean μ and standard deviation σ of the Gaussian distribution is calculated as follows. If W_i is the weight vector

$$\mu = \frac{1}{M}\sum_{i=1}^{M} W_i \quad (7)$$

$$\sigma = \frac{1}{M}\sum_{i=1}^{M} (W_i - \mu) \quad (8)$$

There are different probability distributions and of which some can be fitted more closely to the observed frequency of the data than others. Difference distributions were fitted as test onto the depth weight vector and the Gaussian distribution seemed to be more effective when the data distribution is concerned.

Gaussian pdf is represented by the function

$$f(x,u,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad (9)$$

Corresponding CDF is given by

$$F(x) = \sum_{xi \le x} f(X = x_i) \quad (10)$$

Where F(x) is called the Cumulative Data Feature Vector (CDFV).

Score Fusion

Next the error score is estimated using all the multiple representations separately. For processing 2D DCT representation, 2D-PCA is used. 1D-PCA was also experimented for the 2D representations but 2D-PCA gave better result for 2D representations. Error values are calculated for each data representations and all this error values are combined as a single error value using the linear expression as given in equation 11.

Error= W*Texture_Error_DCT_Norm+ (1-W)*Depth_Error_DCT_Norm (11)

$$\text{Texture\_Error\_DCT\_Norm} = \frac{\text{Texture\_ErrorDCT} - \min(\text{Texture\_ErrorDCT})}{\max(\text{Texture\_ErrorDCT}) - \min(\text{Texture\_ErrorDCT})} \quad (12)$$

$$\text{Depth\_Error\_DCT\_Norm} = \frac{\text{Depth\_ErrorDCT} - \min(\text{Depth\_ErrorDCT})}{\max(\text{Depth\_ErrorDCT}) - \min(\text{Depth\_ErrorDCT})} \quad (13)$$

By trial and error, the optimum value for W can be approximately obtained in the range -1 and +1.

PERFORMANCE ANALYSIS

Results of Fusion Scheme analysis

The Tables 1 shown summarize the results obtained by using fusion scheme. Training images were varied from T1 to T4. Using texture information alone Face recognition accuracy (FRA) of a minimum 53% and a maximum of 73% is obtained, while for depth information it is 44% and

69%. Fusion of texture and depth feature vectors significantly increased the accuracy along different orientations. The table here is showing average accuracy over all the orientations rounded to next higher integer. Testing is done over 1300 samples with different orientations and expressions. Figure 10 shows the equivalent graphical representation of Table 1.

Table 1: Results Of Fusion Scheme With Different Number Of Training Images

Scheme	T1	T2	T3	T4
	FRA%	FRA%	FRA%	FRA%
Texture Feature alone	53.00	60.00	61.00	73.00
Depth Feature alone	44.00	54.00	55.00	69.00
Fusion of Texture and Depth features	62.00	69.00	76.00	82.00

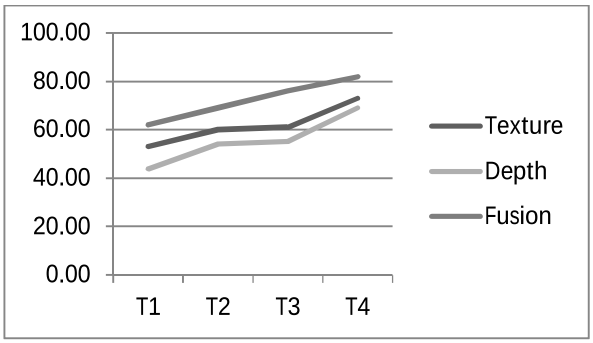


Figure 10. Graph showing improvement in accuracy with increased number of training images

Reliability Test of the Face recognition system with Fusion scheme

The algorithm is tested for robustness against the chances of wrong user authentication (FAR) and the denial of authentication to enrolled users (TRR). The fusion scheme improves the rejection of unauthorized access of a user not enrolled in the database. Table 2 summarizes the averaged results of TAR, TRR, FRR and FAR analysis done along orientations along different axis.

Table 2: Reliability Test Results

Scheme	True Acceptance Rate	True Rejection Rate	False Rejection Rate	False Acceptance Rate
	TAR %	TRR %	FRR %	FAR %
Texture Feature	82.67	17.33	79.33	20.67
Depth Feature	82.00	18.00	64.67	35.33
Fusion of Depth and Texture	82.67	17.33	82.67	17.33

It seems that the fusion scheme has improved the false rejection rate of individual scheme and has reduced the false acceptance rate. Here only a small reduction in the TAR is observed for fusion scheme which can be ignored.

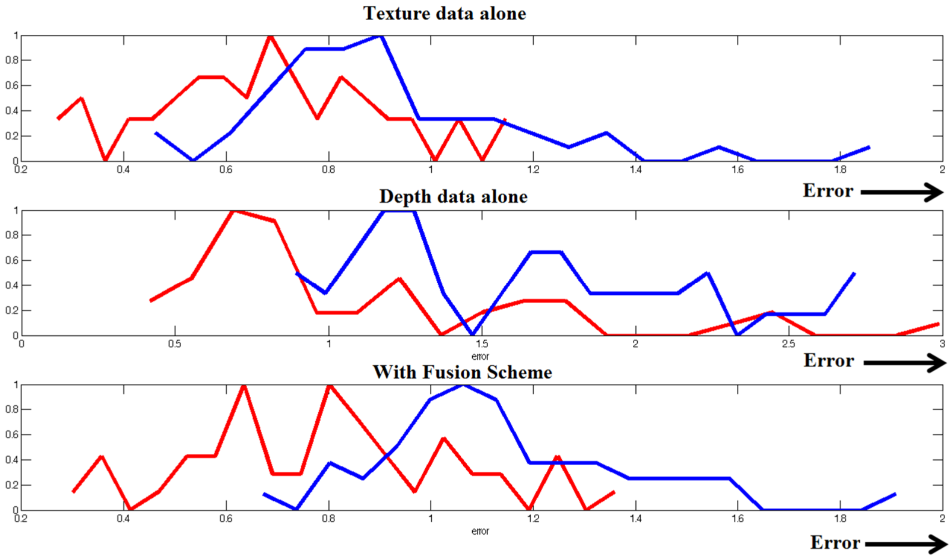


Figure 11: Error Distribution

From Fig. 11, it is clear that with fusion scheme the error distribution is sufficiently separated when compared to the individual error distribution. Red line shows the error value distribution generated when the testing of a subject within the database is done and the blue line shows the error value distribution generated when the test subject is outside the database.

Computation time

Testing of algorithm is done on 3GHz, Core I-5 processor; the average testing time for a single sample is approximately 480ms, 650ms, 1000ms and 1150ms for 1, 2, 3 and 4 training images respectively. That is 480ms for testing and validating a single user. This will again increase as the number of subjects increase. By using down sampling and abstracting the data representations computational time can be further reduced. But with a dedicated system, the testing time can be further reduced to microseconds.

CONCLUSIONS

The fusion algorithm is tested on unregistered 2D Faces with orientations along X, Y and Z axis with facial expressions. The algorithm gives a maximum accuracy of 82%, when tested together over 1300 samples. The experimental results show that the features can be effectively extracted from 2D depth data using statistical representation of spectral information and texture data using spectral transformation. Here fusion experiments were conducted at score level. There is ample scope for further improvement using more fusion schemes at the representation level and at spectral level. The method can also be implemented in real time systems since the processing time required is less.

REFERENCES

http://www.frav.es/databases/FRAV3D

Alexander M. Bronstein; Michael M. Bronstein and Ron Kimmel. Expression-invariant 3D face recognition. In *Proc. International Conference on Audio- and Video-based Biometric Person Authentication*, volume 2688 *of Lecture Notes in Computer Science*, pages 62-70, Guildford, *UK, 2003.*

C. Beumier, 3D face recognition *In IEEE Int. Conf. on Computational Intelligence for Homeland Security and Personal Safety (CIHSPS2004),* Venice, Italy, 21-22 July 2004.

3D Face Recognition using Mapped Depth Images .Gang Pan; Shi Han; Zhaohui Wu; Yueming Wang, Proceedings of the 2005 IEEEComputer Society Conference on CVPR (CVPR'05)Workshops- Volume- 03 p. 175

A Method of 3D Face Recognition Based on Principal Component Analysis Algorithm, Xue Yuan;Jianming Lu;Takashi Yahagi IEEE International Symposium on Circuits and Systems, 2005. 23-26 May 2005 Page(s): 3211 - 3214 Vol. 4

3D Face Recognition Using 3D Alignment for PCA,Trina Russ; Chris Boehnen ;Tanya Peters, Proceedings of the 2006 IEEE Computer Society Conference on CVPR (CVPR'06) Volume 2, 2006 Pages: 1391 – 1398

Automatic 3D Face Detection, Normalization and Recognition, Ajmal Mian;Mohammed Bennamoun;Robyn Owens , Proceedings of the Third International Symposium on 3DPVT (3DPVT'06) June 2006, page(s): 735-742

Gender Recognition Using PCA and DCT of Face Images, Ondrej Smirg, , Jan Mikulka, , Marcos Faundez-Zanuy, Marco Grassi, Jiri Mekyska , Advances in Computational Intelligence , Lecture Notes in Computer Science Volume 6692, 2011, pp 220-227

Pose Normalization for Local Appearance-Based Face Recognition, Hua Gao, Hazım Kemal Ekenel, Rainer Stiefelhagen , Advances in Biometrics, Lecture Notes in Computer Science Volume 5558, 2009, pp 32-41

3D Face Recognition Method Using 2DPCA-Euclidean Distance Classification, Mohammad Naser-Moghaddasi Yashar Taghizadegan, Hassan Ghassemian. ACEEE International Journal on Control System and Instrumentation3(1):5 (February 2012)

3D Face Recognition Using Modified PCA Methods, Omid Gervei, Ahmad Ayatollahi, Navid Gervei, World Academy of Science, Engineering & Technology;Mar2010, Issue 39, p264

M. Turk and A. Pentland, "Eigenfaces for recognition", J. Cognitive Neuroscience , 1991, 3(1), pp. 71 - 86

Jain, A.K. ; Duin, R.P.W. ; Jianchang Mao , Statistical pattern recognition: a review, Pattern Analysis and Machine Intelligence, IEEE Transactions on (Volume:22 , Issue: 1) Jan 2000 pp: 4 - 37