Modified PCA technique to Recognize Facial Expressions

[1]Mr. Shubham Singhal , [2]Dr. Sukhwinder Singh

[1]E&EC Engineering ,PEC University Of Technology, shubhamsinghal2908@gmail.com
[2]E&EC Engineering ,PEC University Of Technology, mfs1980@rediffmail.com

Abstract: Facial expression recognition plays a major role in pattern recognition and image processing. For the face expression recognition three phases are used face detection, feature extraction and classification. In this paper a robust face recognition system is developed and tested. Here 2D Discrete Cosine Transform (DCT) is exploited for feature extraction and certain normalization techniques are invoked that increase its robustness to variation in facial geometry and illumination. The DCT coefficients of face images are truncated in Gaussian or exponential way. Then without doing inverse DCT Principal Component Analysis (PCA) is applied directly for dimensionality reduction. Lastly face recognition task is performed Euclidean distance measurement.

Keywords: Face expression recognition, DCT, Gabor, Wavelet.

I. INTRODUCTION

Expression are a fundamental way to express human emotions and an effective method of non verbal communication [1].psychologist Mehrabian indicated that the verbal part (i.e. spoken words) of a message contributes only for 7 percent to the effect of the message as a whole, the vocal part (e.g. voice intonation) contributes for 38 percent, while facial expression of the speaker contributes for 55 percent to effect of the spoken.

message [2]. With the development of artificial intelligence and pattern recognition, people pay more and more attention to facial expression recognition which is an important technology of intelligent human-interactive interface. Facial expression analysis goes well back in to the nineteenth century [3]. Facial expression analysis has wide range of applications in areas such as video conferencing, user profiling, image retrieval, psychological area, face animation etc [3,4]. Face recognition has become one of most active research areas of pattern recognition. The face recognition methods can be simply classified in to three categories: holistic (global) feature based matching method, local feature based matching method and hybrid matching method [5].for holistic (global) feature based matching method, the whole face region is used as raw input to the recognition system, like principle component analysis (PCA) projection method and an independent Gabor features (IGF) method were applied to face recognition. For local feature based matching method, local features such as eyes, nose, and mouth are first extracted and then their locations and local statistics (geometric and/or appearance) are fed in to a structural classifier. Geometrical features method and elastic bunch graph matching (EBGM) method belong to this category. For hybrid matching method both global and local features are used for the classification.

The general approach to facial expression recognition consists of five steps [8]

1. Image acquisition: - Images used for facial expression recognition are static images and image sequences. Ideally a face acquisition stage features on automatic face detector that allows locating faces in complex scenes with cluttered background.
2. Preprocessing: - Image preprocessing often takes the form of signal conditioning together with segmentation, location or tracking of the face or its parts.

3. Feature Extraction: - Feature Extraction methods can be categorized according to whether they focus on motion or deformation of faces and facial features, respectively. Whether they act locally or holistically.

4. Classification: - Expression categorization is performed by classifiers. Covering parametric as well as non-parametric techniques has been applied to the automatic expression recognition problem.

5. Post processing: - it aims to improve recognition Accuracy by exploiting domain knowledge to correct classification errors.

II. FEATURE EXTRACTION TECHNIQUES

2.1 Discrete Cosine Transform

The Discrete Cosine Transform (DCT) is an invertible linear transform that the data points in terms of sum of cosine

functions oscillate at different frequencies. DCT generates the coefficients from which it is possible to restore back the transformed signal to the original signal by applying the inverse DCT. The 2D-DCT used as the DCT-II, is shown in equations (1) and (2).

Given an input image *f(x, y)* of size *m* x *n*, the 2D *m* × *n* DCT is defined as follows:

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} f(x,y)\cos\frac{\pi(2x+1)u}{2m}\cos\frac{\pi(2y+1)v}{2n} \quad (1)$$

The variables *m* and *n* are coordinates of the space domain and *u* and *v* are the coordinates of frequency domain.

$$\text{Where, } \alpha(u) = \sqrt{\frac{1}{m}} \text{ for } u=0,$$

$$= \sqrt{\frac{2}{m}} \text{ for } u=1,2\ldots\ldots..m-1 \quad (2)$$

$$\text{And } \alpha(v) = \sqrt{\frac{1}{n}} \text{ for } v=0,$$

$$= \sqrt{\frac{2}{n}} \text{ for } v=1,2\ldots\ldots..n-1 \quad (3)$$

In holistic feature extraction method DCT converts high-dimensional face images into low dimensional space in which more significant facial features are maintained. The DCT coefficients are generally divided into three bands as low frequencies, middle frequencies and high frequencies. Low frequencies coefficients are related to illumination variation and smooth regions (like forehead, cheeks etc) of face and high frequencies coefficients represent noise as well as small variations (like edge and details) of face image. The middle frequencies coefficients contain useful information of basic structure of the image which is more suitable candidate for recognition [14]. Hence we can‟t just discard the low frequency components to compensate illumination variations if the image is not so much affected by lighting conditions. Similarly we can‟t just truncate the high frequency coefficients to remove noise as they are responsible for details and edge of the image. In this method we consider all the aspects and accordingly design a mask. The coefficients are modified as exponential or as Gaussian way giving equal weight to the same frequency coefficients i.e. to give same weight to the coefficients of respective slant line. Here we give emphasis to the edge and detail variations of the image but not to the smooth regions and brightness of the image .The accurate reconstruction is not required for face recognition. In the next step PCA has been applied for feature extraction directly to the modified DCT coefficients. In [13], authors proof that as the DCT is an orthogonal transformation, PCA can be directly implemented in the DCT domain.

2.2 Principal Component Analysis

Principal Component Analysis is a standard technique used in statistical pattern recognition and signal processing for data reduction and Feature extraction. Principal Component Analysis (PCA) is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The purpose of PCA is to reduce the large dimensionality of the data space (observed variables) to the smaller intrinsic dimensionality of feature space (independent variables), which are needed to describe the data economically. This is the case when there is a strong correlation between observed variables. The first principal component is the linear combination of the original dimensions that has the maximum variance; the n-th principal component is the linear combination with the highest variance, subject to being orthogonal to the n -1 first principal components.

For example, face image from the database with size 112×92 can be considered as a vector of dimension 10,304, or equivalently a point in a 10,304 dimensional space. An ensemble of images maps to a collection of points in this huge space. Images of faces, being similar in overall configuration, will not be randomly distributed in this huge image space and thus can be described by a relatively low dimensional subspace. The main idea of the principle component is to find the vectors that best account for the distribution of face images within the entire image space. These vectors define the subspace of face images, which we call "face space".

PCA is an information theory approach of coding and decoding face images may give insight into the information content of face images, emphasizing the significant local and global "features". Such features may or may not be directly related to face features such as eyes, nose, lips, and hair. In the language of information theory, we want to extract the relevant information in a face image, encode it as efficiently as possible, and compare one face encoding with a database of models encoded similarly. A simple approach to extracting the information contained in an image of face is to somehow capture the variation in a collection of images, independent of any judgment of features, and use this information to encode and compare individual face images.

These eigenvectors can be thought of as a set of features that together characterize the variation between face images. Each image location contributes more or less of each eigen vector, so that we can display the eigenvector as a sort of ghostly face which we call an ***eigenface***.

Each individual face can be represented exactly in terms of a linear combination of the eigenfaces. Each face can also be approximated using only the "best" eigenfaces-those that have the largest eigenvalues and which therefore account for the most variance within the set of face images. The best M eigenfaces span an M-Dimensional subspace- "face space" – of all possible images.

Consider a image of size N x N which can also be converted into one dimension of size N^2. In which the length of the each dimension vector is N^2[7]. Let us consider that there are Y images of N x N size which can be represented as X_1, X_2,........X_n, then the mean of the data set is

$$y = \frac{1}{Y}\sum_{i=1}^{n} Xi \qquad (4)$$

Then the mean image is subtracted from the each image of the set, so as to equalize the data

$$K_i = X_i\text{-}n \qquad (5)$$

Then a matrix is formed by concatinating all mean images.

$$F=[K_1,K_2,......K_n] \qquad (6)$$

A covariance matrix is formed U= FF^T having dimensions N^2 x N^2 , which then produces eigenvectors and eigenvalues. The eigen vectors are :

$$FF^T J_i=\lambda_i J_i \qquad (7)$$

Which then also be written as

$$F^T FFJ_i = F(\lambda_i J_i) \qquad (8)$$
$$F^T F(FJ_i) = \lambda_i (FJ_i) \qquad (9)$$

FJ_i is the eigen vector denoted by U_i and λ_i is the eigen value. U_i represent the faces which look hazy and are called eigenfaces[5]. The eigenfaces which have the large eigenvalues account for the most variance of the data set. Each face image is now projected on this face space using

$$\omega_r = U^T (X_r - n) \ , \ r = 1,2,......n \quad (10)$$

where $(X_r - n)$ represents the mean centered image . therefore above equation can be used for finding the projection of each image

During recognition process the test image X is projected onto the face space , which obtain the vector ω

$$\omega = U^T (X - n) \qquad (11)$$

The Euclidean Distance ϵ Classifier is used to calculate the distance between the projected test image and each of the projected images during the training

$$\epsilon_r^2 = \ ||\omega - \omega_r||^2 , \ r = 1,2......n \quad (12)$$

A threshold is set up to classify the face

$$\Theta = \frac{1}{2} \max_{t,r} ||\omega_t - \omega_r|| \quad ,j,r=1,2......n \quad (13)$$

III. PROPOSED METHOD

3.1 Light Compensation and Normalization

In this step first the file is read. The height and width of the image are stored in variables 'H' and 'W' respectively. The red green and blue component of the image are stored in the variables 'R', 'G' and 'B' respectively. The image is converted into YcbCr format and is stored in the variable 'YcbCr'. The difference of the red component of this YcbCr image is stored in the variable 'Y'. The minimum and maximum values of 'Y' are stored in the variables 'minY' and 'maxY' and these variables are used to normalize an image to a scale of 255 and the normalized image is stored in the variable 'Y'.A variable 'T' is initialized to 1 and if the average values of the Y is less than 64 then it infers that the image is dark, so to increase the brightness of the image the value of T is increased to 1.4 and if the average value of Y is greater than 192,then it infers that the image is very bright, so to reduce the brightness of image, the value of T is reduced to 0.6.

If the value of 'T' is not equal to 1 then the differences of red component and green component are raised to the power of T**.**

Now a variable 'C' is initialized in which the values of all the red, green and blue components in the detected region are assigned to zero. Now red, green and blue component of the image C are initialized to RI, GI and B.

3.2 Extracting Skin

Image is now converted to yCbCr format because light has least to do with Cb and Cr components. We choose Cr component for skin region detection. First of all we create an image array initialized with 0 of same dimension as of our image .

By our experiment we found that value of Cr component for normal skin in normal conditions ranges from 10 to 45. So we search for co-ordinates where value of Cr is greater than 10 but less than 45 and its row number are stored in SkinIndexRow and column number are stored in SkinIndexCol. Next the co-ordinates(pixels) having value of Cr in between 10 and 45 are turned white(1). One of figure shown below

Fig. 1 Image indicating skin in white

3.3 Remove noise

As there may be some stray white dots outside face region or even black dots in face region. We need to smooth this image. We again create an image array initialized with black(0) and of same height and width. We compute a local sum of 5X5 block and if the sum is greater than 12 that pixel is turned white. This step is basically remove noise from skin region. So get an image of skin with noise removed as shown below.

3.4Algorithm for Face Recognition

- Acquire the face images of the training set and compute the average face image.
- Calculate the deviation of each image from the mean face image, this constitutes the set of mean centered images.
- Obtain the eigenfaces corresponding to this training set, this constitutes the face space
- Now find the projection of the mean-centered images on the face space
- Calculate the Euclidean distance between projected test image and each projected training image
- The threshold value is calculated
- If the minimum Euclidean distance is below the specified threshold, the corresponding test image is considered to be in the database and closest matching image is displayed

Fig. 2. Mean Face Image

Fig 3. Eigen Faces

3.5 Facial expression Detection

3.5.1.Training Process

- Read all the faces of a person in the training database
- Use DCT technique to reduce the data
- Using PCA on the training set calculate eigenvalues
- Calculate eigenvectors of the covariance matrix
- Obtain the corresponding eigenfaces and the projection of the training images onto the face space

3.5.2.Testing Process

- Read all the faces of the person in the test folder
- Project test image onto the face space
- Calculate the Euclidean distance of projected test image from all the projected train images
- The train image with minimum value of the euclidean distance is taken as the face image with closest match to test image.

IV. EXPERIMENTS AND RESULTS

4.1Database

(1) Number of individual : 20
(2) Total number of Training Images : 50
(3) Total number of Testing images : 50
(4) Image Lighting Variation is very little

Fig 4. Output of Various Facial Expressions

TABLE I. Recognition Rates of Various Facial Expressions

Facial Expression	Recognition Rate using Modified PCA
Happy	62.5
Disgust	75
Neutral	87.5
Sad	87.5
Anger	62.5

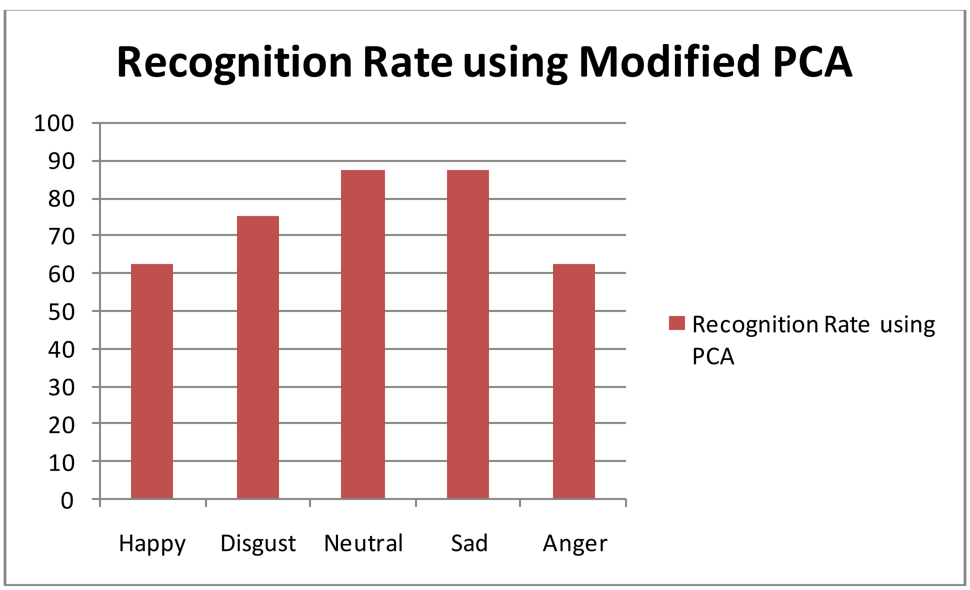

Fig.5 Bar Graph Showing Recognition Rates of Various Facial Expressions

TABLE II. Comparison of results with different techniques

Expressions	DCT	Gabor Filter	Wavelet Transform	Modified PCA
happy	63	55	70	62.5
disgust	65	65	68	75
sad	73	82	71	87.5
neutral	80	80	71	87.5
angry	70	60	70	62.5
average	**70.2**	**68.4**	**70**	**75**

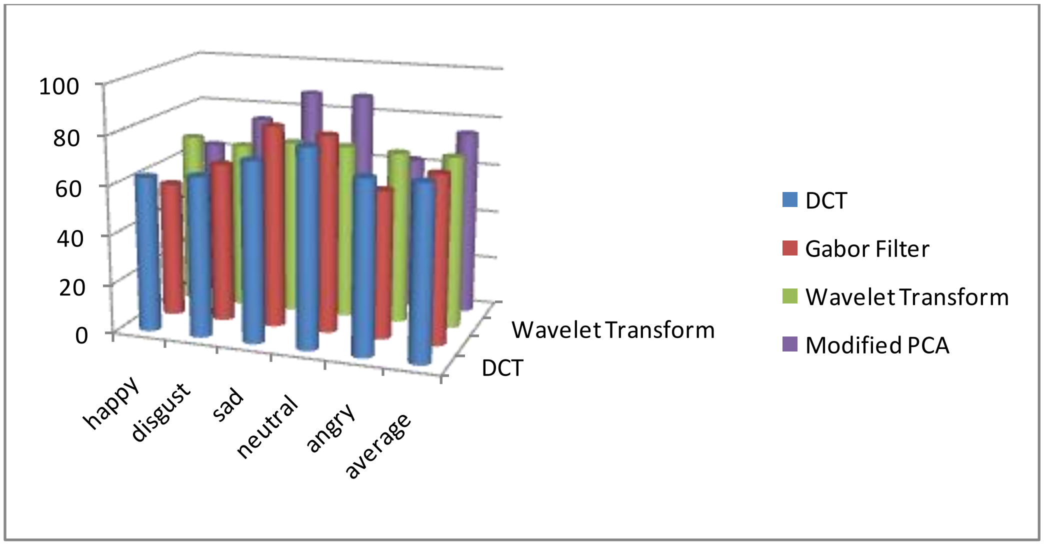

Fig.6. Showing comparison with different techniques

V. CONCLUSION

In this paper facial expression recognition system has been implemented . the experiments shows the recognition rate of 75 % has been obtained using Modified PCA

VI. REFRENCES

[1] Liu Xiao-min, Tan Hua-chun, Zhang Yu-jin."New Research Advances in Facial Expression Recognition".11(10):pp.1359-1368, 2006.
[2] Mehrabian A." Communication without words [J]".Psychology Today.2 (4):pp.53-56, 1968.
[3] B.Fasel and J.Luettin,"Automatic Facial Expression Analysis: A Survey" pattern recognition, Vol.36, pp.259-275, 2003.
[4] M.Pantic, L.Rothkrantz, toward an affect-sensitive multimodal human-computer interaction, proceedings of the IEEE 91(9), pp.1370-1390, 2003.
[5] R.Chellappa, C.L. Wilson and S.Sirohey,"Human and machine recognition of faces: A survey", proceedings of the IEEE, Vol.83, pp.705-740, 1995.
[6] C.Darwin.The expression of emotions in man and animals. John Murray, reprinted by university of Chicago, 1965.
[7] P. Ekmin, W.V. Friesen."Constants across cultures in the face and emotion", Journal of personality Social Psychology, 17(2), pp.124-129, 1971.
[8] C.C.Chibelushi, F.Bourel,"Facial Expression recognition: a brief tutorial overview", on-line Compendium of computer Vision, 2003.

[9] BIN JIANG, GUO-SHENG YANG, HUAN-LONG ZHANG "COMPARATIVE STUDY OF DIMENSION REDUCTION AND RECOGNITION ALGORITHMS OF DCT AND 2D PCA", Proceeding Of The Seventh International Conference On Machine Learning And Cybernetics, Kunming, IEEE Pp. 407-410,12-15 July 2008

[10] Seyed Mehdi Lajevardi and Margaret Lech "Average Gabor Filter Features for Facial Expression Recognition", Digital Image Computing: Techniques and applications IEEE pp. 71-76, 2008.

[11] G. Donato, M. S. Bartlett, J. C. Hager, "Classifying Facial Actions", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 21, 1999, pp.974-989

[12] C. Liu, H. Wechsler, "Independent Component Analysis of Gabor Features for Face recognition", IEEE Trans. Neural Networks, Vol. 14, 2003, pp.919-928.

[13] C. Liu, H. Wechsler, "Independent Component Analysis of Gabor Features for Face recognition", IEEE Trans. Neural Networks, Vol. 14, 2003, pp.919-928.

[14] J. G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression", IEEE Trans. Acoustic, speech and signal processing, Vol. 36, 1988, pp.1169-1179.

[15] Fasel, J. Luettin, "Automatic facial expression analysis: a survey", Pattern Recognition, Vol. 36, 2003, pp.259-275.

[16] Liu, H. Wechsler, "Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition", IEEE Trans. Image Processing, Vol. 11, 2002, pp.467-476

[17] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between Geometry Based and Gabor Wavelets Based Facial Expression Recognition Using Multi-Layer Perceptron," Proc. Third IEEE Conf. Face and Gesture Recognition, pp. 454-459, Nara, Japan, Apr. 1998.

[18] N. Shams, I. Hosseini, M. Sadri, and E. Azarnasab, "Low cost fpgabased highly accurate face recognition system using combined wavelets with subspace methods," pp. 2077–2080, (2006).

[19] C. C. LIU, D. Q. Dai, and H. Yan, "Local Discriminant Wavelet Packet Coordinates for Face Recognition, "Journal of Machine learning Research, pp. 1165–1195, May (2007).

[20] Zheng, D., Zhao,Y., Wang, J.,(2004a) Features Extraction Using A Gabor Filter Family, Proceedings of the sixth LASTED International conference, Signal and Image processing, Hawaii, USA.

[21] T. S. Lee, "Image Representation Using 2D Gabor Wavelets", IEEE Trans. Pattern Analysis and Machine Intelligence,Vol.18,1996,pp.959-97

