An Illumination Invariant Face Recognition by Selection of DCT Coefficients

A.Thamizharasi *radhatamil1@rediffmail.com*
Research Scholar,
Manonmaniam Sundaranar University, Abhishekapatti,
Tirunelveli -627012, Tamilnadu, India

Dr Jayasudha J.S. *jayasudhajs@gmail.com*
Professor,Department of Computer Science & Engineering,
SCT College of Engineering,
Pappanamcode, Trivandrum, Kerala, India

Abstract

The face recognition is nowadays popular in social networks and smart phones. The face recognition is more difficult for poor illumination images. The objective of the work is to create an illumination invariant face recognition system using 2D Discrete Cosine Transform and Contrast Limited Adaptive Histogram Equalization (CLAHE). Contrast Limited Adaptive Histogram Equalization is used for enhancing the poor contrast medical images. The proposed method selects 75% to 100% DCT coefficients and set the high frequency to zero. It resizes the image based on the selection percentage, and then inversed DCT is applied. Then, CLAHE is applied to adjust the contrast. The resized images reduce the computational complexity. The image obtained is illumination invariant face image and termed as 'En-DCT' image. The fisher face subspace method is applied on the 'En-DCT' image to extract the features. The matching face image is obtained using cosine similarity. The face recognition accuracy is tested on AR database. The face recognition is tested with 75% to 100% DCT coefficients and finds the best range. The performance measures recognition rate, 1% FAR (False Acceptance Rate) and Equal Error Rate (EER) are computed. The high recognition rate results prove that the proposed method is an efficient method for illumination invariant face recognition.

Keywords: Face Recognition, 2D DCT, CLAHE, DCT Coefficients Selection, AR, Recognition Rate, Fisher Face, Cosine Similarity.

1. INTRODUCTION

Illumination is considered one of the most difficult tasks for face recognition. Variations caused by pose, expression, occlusion or illumination is highly nonlinear, and making the detection task extremely complex [1]. Here, the proposed work is to normalize illumination. Illumination is a very important problem in face recognition. Research has shown that for a face image, the variability caused by illumination changes even exceeds the variability caused by identity changes [2]. As an example, Fig 1 shows the face images of the same subject under two different illuminations from Yale Face Database. As is evident in Fig 1, the same person, with the same facial expression, can appear strikingly different when light source direction and viewpoint vary. These variations are made even greater by additional factors such as facial expression, perspiration, hair styles, cosmetics, and even changes due to aging.

FIGURE 1: Examples of the same subject seen under different illuminations.

2. RELATED WORKS

The discrete cosine transform (DCT) [3] is a technique for converting a signal into elementary frequency components. It is widely used in image compression. DCT-based image compression relies on two techniques to reduce the data required to represent the image. The first is quantization of the image's DCT coefficients; the second is entropy coding of the quantized coefficients. Quantization is the process of reducing the number of possible values of a quantity, thereby reducing the number of bits needed to represent it. Entropy coding is a technique for representing the quantized data as compactly as possible.

DCT closely approximates KLT in the sense of information packing; it's a very suitable alternative for compact data representation. DCT has been used as a feature extraction step in various studies on face recognition. Up to now, either DCT features have been used in a holistic appearance-based sense [4], or local appearance-based sense which ignores spatial information during the classification step.

A local appearance based face recognition algorithm extracts local information is extracted using block-based discrete cosine transform [5]. The local features are combined both at the feature level and at the decision level. In feature fusion, the DCT coefficients obtained from each block are concatenated to construct the feature vector which is used by the classifier. In decision fusion, classification is done separately on each block and later, the individual classification results are combined.

Xiao-Yuan Jing [6], proposed Discrete Cosine Transform (DCT) to compensate for illumination variations in the logarithm domain. Since illumination variations mainly lie in the low-frequency band, an appropriate number of DCT coefficients are truncated to minimize variations under different lighting conditions.

The recognition rate in ORL database using Eigenfaces [14] approach is 90%. The recognition rate in ORL database using fisherface [15] approach is 82.5%. The recognition rate in ORL database using Direct LDA [16] approach is 90%. The recognition rate in ORL database using discriminant wavelet face method [17] is 94.5%. K Manikantan et. al. [18] proposed a novel Block-Based Discrete Cosine Transform (BBDCT) for feature extraction and a Binary Particle Swarm Optimization (BPSO)-based feature selection algorithm is used to search the feature vector space for the optimal feature subset. This method is tested in ORL database with nine train numbers per subject and one test number per subject. The recognition accuracy rate is 99%.

3. 2D DISCRETE COSINE TRANSFORM (2D DCT)

Like other transforms, the Discrete Cosine Transform (DCT) attempts to de-correlate the image data. After de-correlation each transform coefficient can be encoded independently without losing compression efficiency. This section describes the DCT and some of its important properties. The principle advantage of image transformation is the removal of redundancy between neighboring pixels [8]. This leads to uncorrelated transform coefficients which can be encoded independently. DCT exhibits excellent de-correlation properties.

The most common DCT definition of a 1-D sequence of length N is

$$C(u) = \alpha(u)\sum_{x=0}^{N-1} f(x)\cos\left[\frac{\pi(2x+1)u}{2N}\right], \quad (1)$$

for u = 0,1,2, , *N* –1 … . Similarly, the inverse transformation is defined as

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)C(u)\cos\left[\frac{\pi(2x+1)u}{2N}\right], \quad (2)$$

for X= 0,1,2, , N-1. In both equations **(1)** and **(2)** α*(u)* is defined as

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & for \quad u = 0 \\ \sqrt{\frac{2}{N}} & for \quad u \neq 0. \end{cases} \quad (3)$$

It is clear from **(1)** that for $u = 0$, $C(u=0) = \sqrt{\frac{1}{N}}\sum_{x=0}^{N-1} f(x)$.]

Thus, the first transform coefficient is the average value of the sample sequence. In literature, this value is referred to as the *DC Coefficient*. All other transform coefficients are called the *AC Coefficient.*

The 2-D DCT is a direct extension of the 1-D case and is given by

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right], \quad (4)$$

for u,v 0,1,2, , N-1 and *α(u)* and *α(v)* are defined in equation (3). The inverse transform is defined as

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \alpha(u)\alpha(v)C(u,v)\cos\left[\frac{\pi(2x+1)u}{2N}\right]\cos\left[\frac{\pi(2y+1)v}{2N}\right],$$ for x,y= 0,1,2, , N-1. [5].

4. CONTRAST LIMITED ADAPTIVE HISTOGRAM EQUALIZATION (CLAHE)

Contrast enhancement techniques are used widely in image processing. One of the most popular automatic procedures is histogram equalization (HE) [13]. The Contrast Limited Adaptive Histogram Equalization (CLAHE) is an improved version of adaptive histogram equalization. Originally it was developed for medical imaging and has proven to be successful for enhancement of low contrast images such as portal films. Pisano et al. proposed contrast limited adaptive histogram equalization for detecting abnormalities in dense mammograms [7]. This approach process small regions of the image (called *tiles)* using histogram specification for each tile individually. Neighboring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in areas of homogeneous intensity, can be limited to avoid amplifying noise.

5. FISHER FACE SUBSPACE METHOD

Peter N Belhumeur had a projection method called fisherfaces which is based on Fisher's Linear Discriminant and produces well separated classes in a low-dimensional subspace, even under severe variations in lighting and facial expressions [11]. Karman Etemad and Rama Chellappa used the LDA of faces which provides a small set of features that carry the most relevant information for classification purposes [10]. The features are obtained through eigenvector analysis of scatter matrices with the objective of maximizing between-class variations and minimizing within-class variations.

Belhumeur states, "Using class specific linear methods for dimensionality reduction and simple classifiers in the reduced feature space, one may get better recognition rates than with either the Linear Subspace method or the Eigenface method [11]". Fisher's linear discriminant analysis (FLD) is an example of a class specific method, in the sense that it tries to "shape" the scatter in order to make it more reliable for classification. Fisherface method [12] selects variable 'W' in such a way that the ratio of the between-class scatter and the within class scatter is maximized.

The fisherface similarity is computed using the cosine similarity measure. The cosine of two vectors can be derived by using the Euclidean dot formula in equation 6.

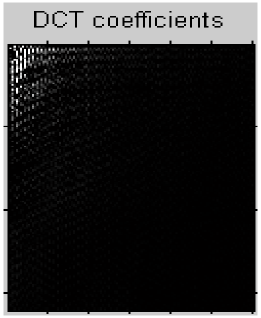
$$a.b = \|a\|\|b\|\cos\theta \tag{6}$$

The cosine similarity Cos(θ) for vectors A and B is represented using a dot product and magnitude as in equation 7.

$$\cos(\theta) = \frac{A.B}{\|A\|\|B\|} = \frac{\sum A_i \quad X \quad B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2} \quad X \quad \sqrt{\sum_{i=1}^{n}(B_i)^2}} \tag{7}$$

6. OBJECTIVE OF WORK

The illumination invariant face image is obtained using the following steps: The input image is color image of size 'm x n x 3' and it is converted into gray scale image of 'm x n'. The log transformation is applied on the gray scale image in order to do the DCT in logarithm domain. The 2D Discrete Cosine Transform is then applied to the logarithm transformed image. The DCT coefficients of size 'm x n' size is now obtained. The energy of the DCT coefficients is concentrated in the upper left corner. The white spots in figure 2a show the energy concentration. The first DCT component is DC component and it carries the illumination and facial information. In the literature the DC component is set to zero and the DCT coefficients are selected based on the compression required. A lot of research work is carried out for selecting the DCT coefficients. The image size in AR database is 160 by 120 and here 'm' is not equal to 'n'. In Figure.2b dark gray area shows the selection of DCT coefficients and light gray area shows the high frequency components set to zero. If there are 'm' rows and 'n' columns, then various percentages of 'm' rows and 'n' columns are chosen.



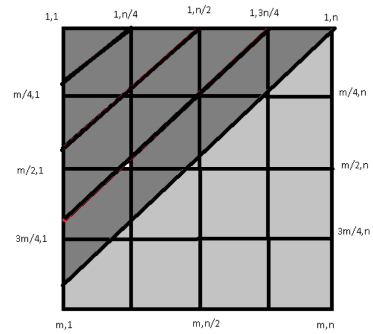


FIGURE 2a: DCT coefficients Coefficients.

FIGURE 2b: Selection of DCT.

TABLE 1: Selection of DCT coefficients by various percentages of original images.

Image rows 'm'	**Image columns 'n'**	**Percent of original image %**
160	120	100
157	118	98
152	114	95
147	110	92
146	109	91
144	108	90
136	102	85
128	96	80
125	94	78
120	90	75

If 'm' rows and 'n' columns are chosen, then it is termed as 'DCT 100'. The various percentages of DCT coefficients selected for testing are 75%, 78%, 80%, 85%, 90%, 91%, 92%, 95%, 98% and 100%. Table 1 shows the Selection of DCT coefficients by various percentages and the corresponding image size. The image is then resized to size based on percentage selection of DCT coefficients. If 75% DCT coefficients is selected then mage size is resized to 120 by 90. The high frequency components are set to zero as shown in light gray area in fig 1. The inverse DCT is then applied. The next step is to apply logarithm transform to enhance the gray levels. The contrast in the image is adjusted by applying CLAHE and logarithm transform. The image obtained is the enhanced image and termed as 'En-DCT' image. En-DCT is the illumination invariant face image used for testing the face recognition accuracy. The above steps are shown as block diagram in figure 3.

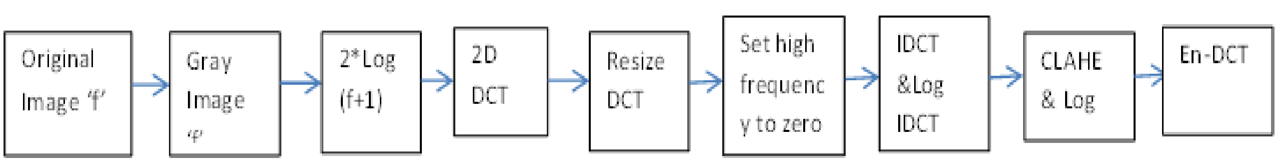


FIGURE 3: Block Diagram of Proposed Method.

7. EXPERIMENTAL WORK

The experimental work is done using AR database. The details of AR database are given below. The AR database (Martinez and Benavente, 1998, 2000) contains images of 100 persons taken in two different sessions. 50 are men and 50 are women [9]. In each session photos were taken in 13 different conditions. They are Neutral, expressions (anger, scream, and smile) ,illumination (right, left and both sides) and occlusions (eye occlusion, eye occlusion with left illumination, eye occlusion with right illumination, mouth occlusion, mouth occlusion with left illumination and mouth occlusion with right illumination).All the images are cropped to an image size of 160 × 120 pixels. There are 26 images for each person and totally there are 2600 images. AR face image subset is created with 14 images of each person with a total of 1400 images. This subset contains images of different illuminations and different expressions and neutral conditions. Out of the 14 images of each person, the first image in neutral condition is used as train image. The testing is done by changing the total number of train images per subject from 2 to 10. Figure 4 shows the sample image in AR database subset. Figure 5 shows the series of output images obtained by using the block diagram shown in figure 3. The first image is input image and the last image is 'En-DCT' image which is an illumination invariant face image.

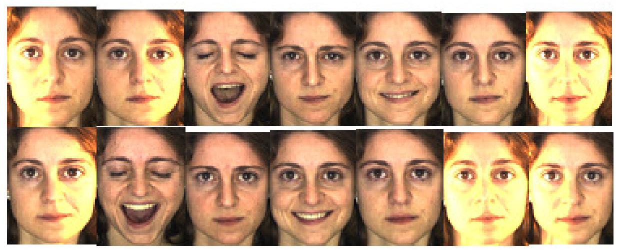


FIGURE 4: Sample images in AR database for a single subject with different illumination and expression.

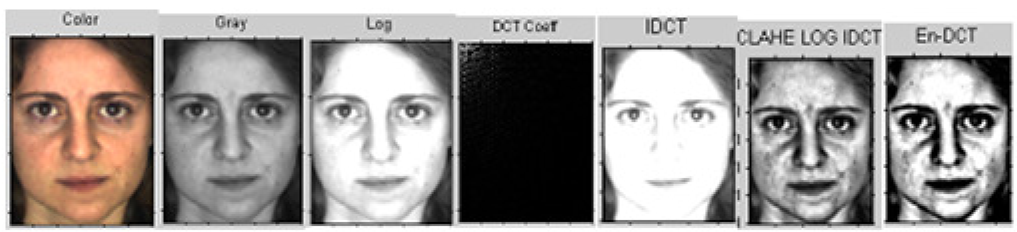


FIGURE 5: Creation of En-DCT from original color image using the proposed method.

8. RESULTS AND DISCUSSION

The testing is done using AR database. There are 100 subjects in AR database. The testing is done by using the five train numbers per subject. If the total number of train numbers per subject is 5, then there will be 9 images per subject for testing. Thus, train database contains 500 images and test database contains 900 images. The train database and test database are created with En-DCT images. The fisher face sub space projection method is used to extract the features from train and test database. The similarity among the test image features and train image features is compared by using cosine similarity. The performance measures face recognition rate, 1% False Acceptance Rate (FAR) and Equal Error Rate (EER) are computed. Table 2 shows the Performance measures of 'En-DCT' on AR database using fisher face sub space with cosine similarity by various percentages of DCT coefficients selection. If 100% selection is chosen i.e. 'm' rows and 'n' columns are chosen then the recognition rate is 91.67% and 1% FAR is 94.56%. However 90% to 100% selection of DCT coefficients shows recognition rate greater than 90% and 1% FAR greater than 93%. If 90% selection is used, then the image size is reduced from 160 by 120 to 144 by 108. If the selection is 75%, the image size is 120 by 90; recognition rate is 82% and 1%FAR is 91.11. The testing is done with 50% selection, the image size is 80 by 60; recognition rate is 70.22% and 1% FAR is 78.78%. The Equal Error rate at 95% selection is 2.87 with image size 152 by 114.

The performance of the proposed method is compared with the following works. Vikas Maheshkar et. al. [19] use 8X 8 DCT Blocks and in these blocks the diagonal coefficients are selected and feature vector is created. The face recognition is tested in ORL database with 80 train images and 40 test images; it results in recognition of 35 images. Sujatha B.M et al [20] had shown the recognition rate of 93.33% in ORL database. The 1% FAR performance of the proposed method is greater than 93.33% when using 78% to 100% DCT coefficients. The performance of the proposed method is better than eigen face, fisher face, direct LDA methods.

TABLE 2. Performance measures of En-DCT on AR database using fisher face with varying DCT coefficients selection.

Performance Measures in %	**En-DCT +Fisher face + Cosine similarity for AR**									
	Percentage Selection of DCT coefficients									
	75	**78**	**80**	**85**	**90**	**91**	**92**	**95**	**98**	**100**
Recognition rate	82.44	87.89	90	91.11	88.78	90.56	91.11	91.56	91.33	91.67
1% FAR	91.11	93.44	93.22	94.33	93.89	93.89	94.33	95	94.44	94.56
EER	4.11	3.99	4.08	3.55	4.21	3.47	3.58	2.87	3.62	3.22

9. CONCLUSION

The face recognition accuracy is affected due to illumination variations. The face recognition accuracy could by increased by first normalizing the illumination of the image. This work is done to create illumination invariant face recognition system by using 2D DCT and CLAHE. 2D DCT is applied to logarithm image and DCT coefficients are extracted. The DCT coefficients are selected by various percentages of original image size. The image size is resized based on the percentage selection. The high frequency components of DCT coefficients are set to zero. The CLAHE is applied to adjust the contrast in the image. The contrast adjusted image is enhanced by logarithm transform. The image obtained is thus enhanced and illumination invariant image. The efficiency of the method is tested on AR database. The fisher face feature extraction is applied and test image is matched to train image using cosine similarity. The selection of 95% DCT coefficients shows recognition rate of 91.56%, 1% FAR is 95% and EER is 2.87.

10. REFERENCES

[1] A. Georghiades, P. Belhumeur, and D. Kriegman, “From few to many: Illumination cone models for face recognition under variable lighting and pose,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 643–660, 2001.

[2] Thomas Heseltine1, Nick Pears, Jim Austin, Zezhi Chen, Face Recognition: A Comparison of Appearance-Based Approaches, Proceedings of 7th Digital Image Computing Techniques and Applications, Dec 2003.

[3] Andrew B. Watson, NASA Ames Research Center, Image Compression Using the Discrete Cosine Transform, Mathematica Journal, 4(1), p. 81-88, 1994.

[4] Z.M.Hafed and M. D. Levine, Face recognition using the discrete cosine transform, International Journal of Computer Vision, 43(3), 2001.

[5] Hazim Kemal Ekenel , Rainer Stiefelhagen , Local Appearance based Face Recognition Using Discrete Cosine Transform (2005), 13th European Signal Processing Conference (EUSIPCO), 2005.

[6] Xiao-Yuan Jing and David Zhang, A Face and Palm print Recognition Approach Based on Discriminant DCT Feature Extraction, IEEE Transactions on Systems, Man and Cybernetics, Vol.34, No.6, 12, 2005.

[7] E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning and S. M. Pizer, “Contrast Limited Adaptive Histogram Equalization Image Processing to Improve the Detection of Simulated Spiculations in Dense Mammograms,” Journal of Digit Imaging, Vol. 11, No. 4, pp. 193-200, 1998.

[8] Ziad M.Hafed and Martin D. Levine, Face Recognition Using the Discrete Cosine Transform, International Journal of Computer Vision 43(3), 167–188, 2001.

[9] Martinez A and Benavente R, “AR Face Database”, Technical Report 24, Computer Vision Centre, 1998.

[10] K.Etemad and R. Chellappa,"Discriminant analysis for face recognition of human face images", Journal of Optical society of America, Vol. 14, No 8, pp. 1724- 1733, 1997.

[11] P.N.Belhumeur, J.P.Hespanha and D.J. Krieg man, “Eigen faces Vs. Fisher faces: Recognition using Class specific linear projection”, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol 19, No. 7, pp. 711-720, 1997.

[12] R. Chellappa, C. Wilson, and S. Sirohey, “Human and Machine Recognition of Faces: A Survey,” Proceedings of IEEE, vol. 83, no. 5, pp. 705- 740, 1995.

[13] Rafael C. Gonzalez and Richard E. Woods, “Digital Image Processing”, Addison- Wesley, 1993.

[14] M. Turk and A. Pentland, “Eigenfaces for recognition,” International Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71–86, 1991.

[15] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. fisherface: Recognition using class specific linear projection,” IEEE Trans. Pattern Anal. Machine Intell., vol. 19, pp. 711–720, July 1997.

[16] H. Yu and J. Yang, “A direct LDA algorithm for high-dimensional data with application to face recognition,” Pattern Recognit., vol. 34, no. 12, pp. 2067–2070, 2001.

[17] J. T. Chien and C. C. Wu, “Discriminant waveletfaces and nearest feature classifiers for face recognition,” IEEE Trans. Pattern Anal. Machine Intell., vol. 24, pp. 1644–1649, Dec. 2002.

[18] K Manikantan, Vaishnavi Govindarajan, V V S Sasi Kiran, S Ramachandran, Face Recognition using Block-Based DCT Feature Extraction, Journal of Advanced Computer Science and Technology, (4) (2012) 266-283.

[19] Vikas Maheshkar , Sushila Kamble , Suneeta Agarwal , and Vinay Kumar Srivastava, “DCT-Based Reduced Face for Face Recognition”, International Journal of Information Technology and Knowledge Management January-June 2012, Volume 5, No. 1, pp. 97-100.

[20] Sujatha B M, K Suresh Babu, K B Raja & Venugopal K R, Hybrid Domain based Face Recognition using DWT, FFT and Compressed CLBP, International Journal of Image processing (IJIP), Volume (9) : Issue (5) : 2015.