ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

AUTOMATIC FACIAL EXPRESSION RECOGNITION USING DISCRETE COSINE TRANSFORM

Rahul Manche and Abhay Kshirsagar

Manuscript Info

Manuscript History:

Received: 13 October 2013
Final Accepted: 28 October 2013
Published Online: November 2013

Key words:
Face recognition, discrete cosine transform, geometric normalization, illumination normalization, feature extraction.

Abstract

In this paper, facial expression recognition (FER) method is explicated which uses the 2-D DCT algorithm to evaluate the expression of the face. This method uses minimal number of steps, which includes normalization techniques, compares the test image, whose expression is to be calculated, with the standard pre-defined images and concludes the expression of the face as that of image in the database with whom the test image produces minimum Euclidean distance. Experiments and the results demonstrated show that this method has good ability in calculating the expression accurately.

Introduction

In human conversation facial expression plays a vital role and a famous apothegm goes like ‘emotions speak more than words’. Automatic FER will help in ameliorating the communication between man & machine. The rise in significance of robots in human being’s life thus can be eased with robot automatically interpreting the expression of his master.

The copious numbers of algorithms have already been proposed to conclude the human emotion.

The expression of face is contributed by various parts of the face like eyes, nose, and lips etc. which are called *facial features*. Paul Ekman and Wallace V. Friesen introduced ‘Facial Action Coding System’(FACS) back in 1978 wherein the face was divided into *Action Units(AU)* including aforesaid features of face [4]. In FACS the muscle movements of these feature-bearing AUs are used to conclude any human facial expression of interest. The FACS algorithm divides the 3-dimensional (3-D) facial model into about 44 AUs which is a complex process.

The recent technology uses 2-D digital images which contain a lot of data. So, the algorithms now need not be that complex as the processing of a 2-D image is much easier. Given that recognizing faces is critical for humans in their everyday activities, automating this process would be very useful in a wide range of applications including security, surveillance, criminal identification, and video compression.

RELATED WORK

Viola-Jones method to detect face uses object detection algorithm to detect objects real-time. Its features are derived from pixels. Pentland et al use Eigenspace method wherein they apply Principal Component Analysis (PCA) on sample of 128 images [10]. Eigenfaces define subspace of sample images i.e. facespaces. Spatio temporal filtering performed to detect faces the filtered image is thresholded in order to further analyze motion blobs and each motion blob represents a human head is then evaluated as a single image. The Potential Net method uses a centre of eyes and centre of mouth to normalize input image [13]. This algorithm applies an integral projection method which synthesizes color and edge information. Potential Net is fitted to normalized image to model the face and its movement with only a condition that face should be devoid of facial hair, glasses etc.

The facial feature extraction method proposed by Yacoob & Davis is done on image sequences real time [16]. The method develops a region tracker for rectangles enclosing facial features. Each rectangle encloses one feature of interest so the flow of computation within region is not contaminated by motions of facial features. In eigenface method face and its feature positions viz. eyes, nose etc are extracted [10]. To detect location of prominent facial feature distance of each feature image from relevant feature space is calculated using FFT and a local energy computation. This extracted position is used to normalize the input image. Kimura and Yachida propose a method in which first step is to compute edge of normalized image [13]. Then image is filtered using Gaussian filter. Potential Net is fitted to each frame of examined facial image sequence. They then compare pattern of deformed net with the pattern extracted from expressionless face and variation of net nodes is further processed. The Kotsia and Pitas propose a method where user manually places candid wise frame model at first frame of image sequence [17]. Tracking system allows candide wireframe to follow facial expression subset of candide grid modes contributing to FACS. Geometrical displacement of candide grid nodes, the difference of each node coordinates at the first and the last frame of facial image sequence used for FER.

Facial expression classification using Template based is wherein facial expression is compared to predefined templates for each expression [18]. The best match decides classification of shown facial action. Spatio temporal method proposed by Essa and Pentland generates spatio temporal templates for six different facial expressions [19]. Euclidean distance between motion energy template and observed motion energy template is used for FER.

PROPOSED METHOD

The method proposed uses the Discrete Cosine Transform's (DCT) property to assign a unique frequency value to a component as per its magnitude. It is proposed to apply DCT to face wherein all the features of face have a peculiar frequency value after DCT is applied to it. Now, only those frequency components are taken which are vital for face recognition and expression recognition.

Normally, the algorithms specify to evaluate *feature vector*, a vector consisting of all frequencies of facial features [1]. But the method we have used over here doesn't need to calculate feature vector. Only one frequency component is enough to conclude the expression of the required face, thus reducing the calculating time significantly.

IMPLEMENTATION STEPS

(A) 2-DIMENSIONAL DISCRETE COSINE TRANSFORM (2-D DCT)

Data compression is essential for both biological and computer signal processing. Data compression is the main feature of the discrete cosine transform. Also, since the DCT is related to the discrete Fourier transform (Rao and Yip, 1990), it can be computed efficiently. It is these two properties of the DCT that we seek for face recognition.

Ahmed, Natarajan and Rao (1974) first introduced the discrete cosine transform (DCT) [2]. The DCT was categorized by scientists later into four slightly different transformations named DCT-I, DCT-II, DCT-III, and DCT-IV.

For an input sequence $x(n)$ the 2-D DCT is given by the following equation

$$\boxed{y(k) = a(k) \sum_{n=0}^{N-1} u(n) \cos\left(\frac{(2n+1)\pi k}{2N}\right)}$$

$0 \le k \le N-1$ …(1a)

Where

$$a(0) = \sqrt{\frac{1}{N}}, \; a(k) = \sqrt{\frac{2}{N}} \quad 1 \le k \le N-1$$

…(1b)

Alternatively, we can think of the sequence $x(n)$ as a vector and the DCT as a transformation matrix applied to this vector to obtain the output $y(k)$. In this case, the DCT transformation matrix, $C = \{c(k,n)\}$, is defined as follows:

$$c(k,n) = \begin{cases} \dfrac{1}{\sqrt{N}}, & k = 0, 0 \le n \le N-1 \\ \sqrt{\dfrac{2}{N}} \cos\left(\dfrac{(2n+1)\pi k}{2N}\right), & 1 \le k \le N-1, \\ & 0 \le n \le N-1 \end{cases}$$

…(2)

Where k and n represent row and column indices respectively.

From the above mentioned equations we observe that for a given sequence DCT converts it into weighted sum of basis cosine sequences.

(B) NORMALIZATION

When the test image is fed for expression recognition it contains a lot of data which is unnecessary for the expression recognition and plays no role in it. So, it is better to remove the extraneous part and concentrate only on the face for further analysis.

The method used is to crop face only to retain facial features which contain all the data needed for expression recognition. The face is cropped in such a way that face contains eyebrows, eyes, nose, and lips. The rest of the face is unused. Hence by doing so solves the problem of geometric normalization wherein some test faces are not of uniform size, thus by cropping them geometric normalization can be further reduced. The pixel size is so adjusted that test image face and database image face both have same number of pixels and thus obviating the need to extract facial features from the DCT matrix.

Next step is illumination normalization. To make all faces, including test images and database images, of uniform intensity they all are converted to grayscale runtime. Thus error due to variation in illumination is eliminated.

DCT is performed on the test image, which is normalized using above steps, and the database image.

(C) EUCLIDEAN DISTANCE COMPUTATION

Many methods which were earlier introduced propose to calculate the Euclidean distance of the DCT matrices of test image and database images.

But the experimental analysis done showed there was no need to extract facial feature vector from the DCT matrix. This paper proposes to just find out the frequency component with least magnitude from the DCT matrix of test image. The least magnitude frequency component from the DCT matrix of database image is calculated. The Euclidean distance is calculated between the least magnitude frequency component of test image and the least magnitude frequency component of all the database images.

(D) FACIAL EXPRESSION CLASSIFICATION

On calculation of Euclidean distances of test image with various database images, the Euclidean distance with least magnitude is used for expression classification. The database image with which the test image produces least Euclidean distance is concluded to be the expression of the test image.

IV. EXPERIMENTAL RESULTS & SEQUENCE OF STEPS

To implement the algorithm proposed in this paper simulating software MATLAB was used on the faces of FEED database. The database used for expression classification is of Carnegie Melon University.

The database consisted of faces of 9 people in various expressions. The database images were 64x64pixel size, grayscale, bmp images.

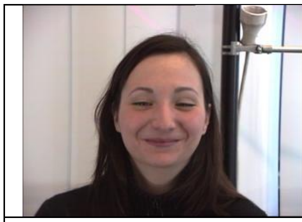
The input test image is read and cropped to remove unwanted data. Since the database image size is 64x64pixels, the cropped face is adjusted to 64x64pixels. Of course some data was lost in this process but it wasn't significant and retained image included all facial features. Thus geometric normalization is achieved.

To perform illumination normalization all geometric normalized images are converted to grayscale. There's no need for illumination normalization to be performed on database images as they were already grayscale.

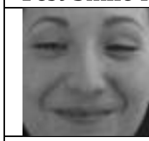
Next step is to calculate DCT of test input image and database images and compute Euclidean distance with least component of DCT matrices of test image and database images.

The database image with which the test image produces least Euclidean distance is concluded to be the expression of the test image.

Fig (I) Results from smile detection

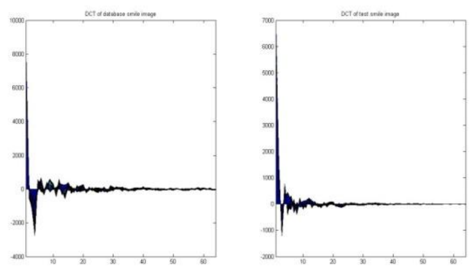


Test Smile face



Normalized Smile face

Corresponding Smile face in Database



Graph showing DCT of test smile image & DCT of corresponding matching smile image in database

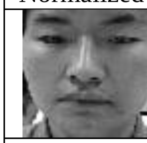
Table (I) Accuracy table for smile detection

	Smile	**Neutral**	**Surprised**
Smile	91.66%	8.33%	-

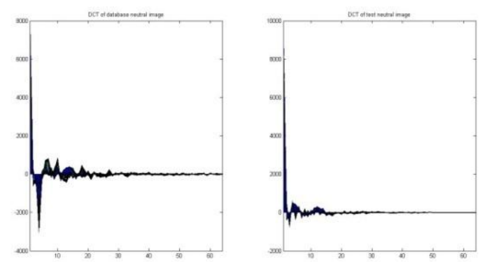
Fig (II) Results from neutral detection

Test Neutral face

Normalized Neutral face



Corresponding Neutral Face in Database



Graph showing DCT of test neutral image & DCT of corresponding matching neutral image in database

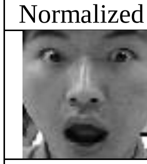
Table (II) Accuracy table for neutral detection

	Smile	**Neutral**	**Surprised**
Neutral	10.29%	82%	7.71%

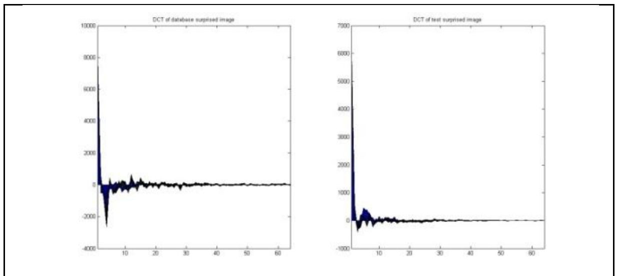
Fig (III) Results from surprise detection

Test Surprised face

Normalized Surprised face



Corresponding Surprise Face in Database



Graph showing DCT of test neutral image & DCT of corresponding matching neutral image in database

Table (III) Accuracy table for Surprise detection

	Smile	Neutral	Surprised
Surprise	6.74%	6.6%	86.66%

V. HIGH NOTE

Extracting facial expression is paramount in human-machine interaction. The application of this method is of very useful, accurate and hassle-free because it is non-intrusive. The various applications include uses in multi-player online games wherein the facial expression determines the next step of the game, uses in Human-Computer Interface (HCI), and also to control music player, where the expression of the face determines the genre of music being played.

Conclusion

An easiest, quick, and holistic approach to classify facial expression was investigated and tested. The approach was based on the discrete cosine transform, and experimental evidence to confirm its usefulness and accuracy was presented. Among the various steps employed face normalization techniques were also incorporated in the face recognition system discussed here.

References

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

[2] Ahmed, N., Natarajan, T., and Rao, K. 1974. “Discrete cosine transform.” *IEEE Trans. on Computers*, 23(1):90–93.

[3] Swarup Ku. Dandpat and Sukadev Meher,

“New Technique for DCT-PCA Based Face Recognition”.

[4] P. Ekman and W. Friesen, “Facial Action Coding System”, Consulting Psychologists Press, 1978.

[5] Nazil Perveen, Shubhrata Gupta and Kesari Verma, “Facial Expression Recognition Using Facial Characteristic Points and Gini Index”, 978-1-4673-0455-9, IEEE 2012.

[6] Nan Liu, Han Wang, “Recognition of Human Faces using Discrete Cosine Transform Filtered Trace Features”, 1–4244–0983–7/07, IEEE 2007.

[7] Keith Anderson and Peter W. McOwan, “A Real-Time Automated System for the Recognition

of Human Facial Expressions”, IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, Vol. 36, No. 1, February 2006.

[8] Shermina J., "Illumination Invariant Face Recognition Using Discrete Cosine Transform and Principal Component Analysis", Proceedings of ICETECT 2011.

[9] R A Patil, Vineet Sahula, A. S. Mandal, "Automatic Recognition of Facial Expressions in

Image Sequences: A Review", 2010 5th International Conference on Industrial and Information Systems, ICIIS 2010, Jul 29 - Aug 01, 2010, India.

[10] Matthew A. Turk and Alex P. Pentland, "Face Recognition Using Eigenfaces", CH2983-5/91/0000/0586, IEEE 1991.

[11] Yonghua Xie, Lokesh Setia, Hans Burkhardt, "Block DCT Vectors Construction for Face Retrieval

Based on Genetic Algorithm", Third International Conference on Natural Computation (ICNC 2007).

[12] B. Fasela and Juergen Luettin, "Automatic Facial Expression Analysis: A Survey", Pattern Recognition 36 (2003) 259 – 275.

[13] Satoshi Kimura, Masahiko Yachida, "Facial Expression Recognition and Its Degree Estimation", 1063-6919/9, IEEE 1997.

[14] Frank Wallhoff: Facial Expressions and Emotion Database *http://www.mmk.ei.tum.de/~waf/fgnet/feedtum.html*, Technische Universität München 2006

[15] CMU AMP Face Expression Database. April 2001.

[16] Y. Yacoob, L.S. Davis, "Recognizing human facial expression from long image sequences using optical flow, IEEE Trans. Pattern Anal. Mach. Intell. 18(6) (1996) 636-642.

[17] I Kotsia and I. Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines", IEEE Trans. On Image Processing, vol. 16, pp. 172-187, Jan 2007.

[18] M. Pantic and L. J. M. Rothkrantz, "Automatic Analysis of facial expressions: The state of the art", IEEE Transactions on Pattern Analysis and machine intelligence, vol. 22, pp. 1424-1445, Dec 2000.

[19] I. A. Essa and A. P. Pentland, "Coding, analysis, interpretation and recognition of facial expressions", IEEE Transactions on Pattern Analysis and machine intelligence, vol. 19, pp. 757-763, July 1997.

[20] www.mathworks.com