ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Classification of Human Emotions for Face and Voice Using KNN Classifier

Meghana B M, K Mahantesh

Dept. Electronics and Communication Engineering, S.J.B. Institute of Technology, Bangalore, India

:

ABSTRACT*:* Face detection is a fundamental task for applications such as face tracking, red-eye removal, face recognition and face expression recognition. The face and emotion detection has experienced increasing attention over years, due to its usefulness in pervasive applications. In order to meet the real-time needs and overcome the high costs and privacy issues, in this paper, we develop a new model for detecting face and emotion using Viola-Jones face detection algorithm and to evaluate the face and emotion detection KNN classifier, MFCC (Mel Frequency Central Co-efficient) feature extraction. Firstly we take image as a input and apply signal processing techniques to convert into time domain signals. To foreshorten face and emotions accurately and efficiently, will use the MFCC along with LBP(Local Binary Pattern) and DOG for feature extraction. On the basis of result from feature extraction the KNN classifier classifies the activity of our input signal. Experimental results demonstrate that the proposed approach is superior to other approaches to recognize.

KEYWORDS: Viola-Jones face detection algorithm, MFCC (Mel Frequency Central Co-efficient), LBP (Local Binary Pattern) and DOG feature extraction.

I. INTRODUCTION

Face detection and emotion selection is the one of the current topic in the security field which provides solution to various challenges. Beside traditional challenges in captured facial images under uncontrolled settings such as varying poses, different lighting and expressions for face recognition and different sound frequencies for emotion recognition. For the any face and emotion detection system database is the most important part for the comparison of the face features and sound Mel frequency components. For database creation features of the face are calculated and these features are store in the database. This database is then use for the evaluation of the face and emotion by using different algorithms.

In this paper we are going implement an efficient method to create face and emotion feature database and then this will be used for face and emotion recognition of the person. For detecting face from the input image we are using Viola-Jones face detection algorithm and to evaluate the face and emotion detection KNN classifier is used. Most of the face recognition (FR) approaches have focused on the use of two dimensional

images. Since FR is still an unsolved problem under the different conditions, such as pose, illumination or database size. The expression of emotions and the recognition of a person's affective state are abilities indispensable for natural human interaction and social integration. The study of emotions has attracted interest of researchers from very diverse areas, ranging from psychology to the applied sciences. Face and emotion features detection is the currently very active area of research in the computer vision field as different kinds of face detection application are currently used such as image database management system, monitoring and surveillance analysis, biomedical image, smart rooms intelligent robots, human computer interfaces and drivers alertness systems.

In last few decades very big amount of research is done in the field of face and emotion features detection. Face recognition and emotion recognition has got more attention and many advance technologies and methods are introduced for this in this period. Many commercial systems for face recognition are now available. New databases have been created and evaluations of recognition techniques using these databases have been carried out. Now, the face recognition has become one of the most active applications of pattern recognition, image analysis and understanding.

Facial recognition plays a vital rule in human computer interaction. A Face recognition system can be either verification or an identification system depending on the context of an application. The verification system authenticates a person's identity by comparing the captured image with his/her own templates stored in the system. It performs a one

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

to one comparison to determine whether the person presenting himself/herself to the system is the person he/she claims to be. An identification system recognizes a person by checking the entire template database for a match. It involves a one to many searches. The system will either make a match or subsequently identify the person or it will fail to make a match.

II. Literature Survey

Viola et.al [1] has made face detection practically feasible in real world applications such as digital cameras and photo organization software. Boosting is a method of finding a highly accurate hypothesis by combining many “weak” hypotheses, each with moderate accuracy V. cc et.al [2] developed a famous and flourishing facial action coding system. The Facial Action Coding System (FACS) identifies the facial muscles that cause changes in the facial expression thus enabling facial expression analysis.

Murthy et.al [3] proposed the enhanced method to recognize the expression of the front view of the face, they used the projected gradient method and developed the graph-preserving sparse non-negative matrix factorization (GSNMF) for extraction of feature verified on different databases. Fasel et al. [4] has proposed the sign-based approach uses a FACS system, encoding action units in order to categorize an expression based on its constituents. This approach assumes no categories, but rather assigns an emotional value to a face using a combination of the key action units that create the expression.

III. Methodology

The main aim of this paper is to implement an efficient method to detect the face and emotion of the person. Work is divided into two parts for the storing the features of the face and the features of the voice of human and second evaluation of the face and emotion of the person using the features database. We use Viola Jones face detection algorithm to create the database for the face detection. The numbers of input images are collected for the creation of database next will create the database for the features of voice with the different emotions of the person. Figure.1 shows the overall implementation of proposed system.

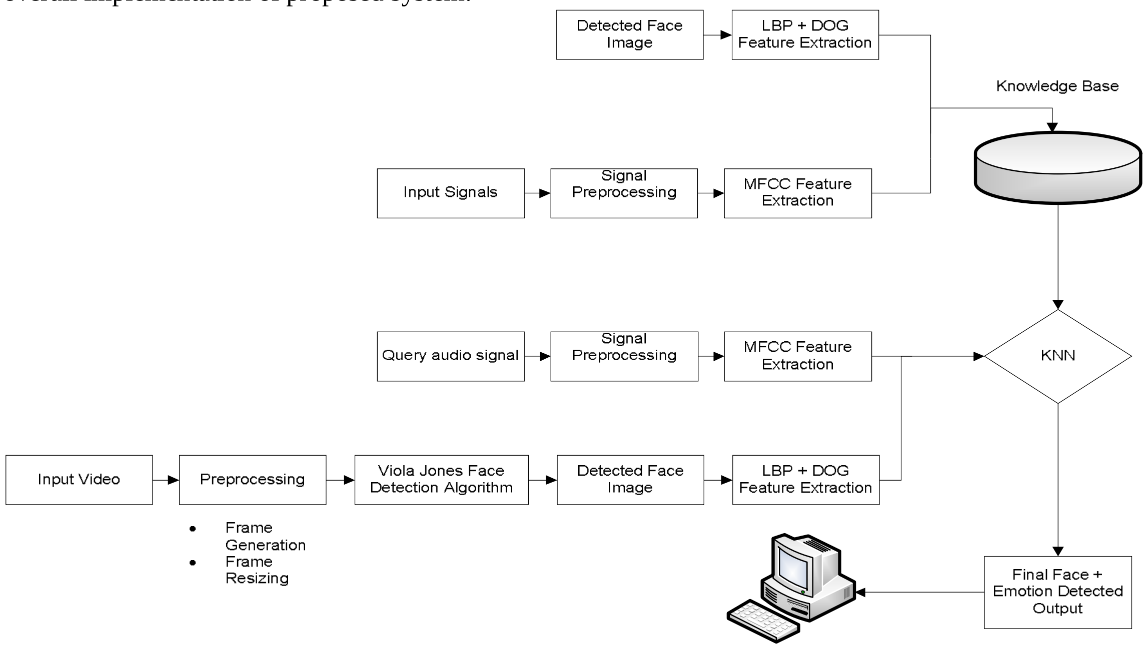


Fig. 1. Block Diagram of Multi Person Tracking System

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

3.1 Viola-Jones Face Detection

Voila Jones is the oldest and most recognized face algorithm available for the face detection from the image. The basic principle of the Viola-Jones algorithm is to scan a sub-window capable of detecting faces across a given input image. The standard image processing approach would be to rescale the input image to different sizes and then run the fixed size detector through these images. This approach turns out to be rather time consuming due to the calculation of the different size images. Contrary to the standard approach Viola-Jones rescale the detector instead of the input image and run the detector many times through the image – each time with a different size. At first one might suspect both approaches to be equally time consuming, but Viola-Jones have devised a scale invariant detector that requires the same number of calculations whatever the size. This detector is constructed using a so-called integral image and some simple rectangular features reminiscent of Haar wavelets.

3.2 Local Binary Pattern (LBP)

Local Binary Pattern (LBP) features have performed very well in various applications, including texture classification and segmentation, image retrieval and surface inspection. The original LBP operator labels the pixels of an image by thresholding the 3-by-3 neighborhood of each pixel with the centre pixel value and considering the result as a binary number. Each face image can be considered as a composition of micro-patterns which can be effectively detected by the LBP operator. To consider the shape information of faces, they divided face images into M small non-overlapping regions $R_0, R_1, \ldots\ldots R_M$.

The LBP histograms extracted from each sub-region are then concatenated into a single, spatially enhanced feature histogram defined as:

$$H_{i,j} = \sum_{x,y} I(f_1(x,y) = i)\ I((x,y)\epsilon R_j \qquad (1)$$

Where i = 0, ... , L-1, j = 0, ..., M-1. The extracted feature histogram describes the local texture and global shape of face images.

3.3 DOG (Difference of Gaussians)

DOG is a feature enhancement algorithm explains the subtraction of one blurred version of an original image from another, less blurred version of the original. In the simple scenarios of gray scale images, with differing standard deviations the blurred images are obtained by convolving the original gray scale images with Gaussian kernels. Blurring an image using a Gaussian kernel suppresses only high-frequency spatial information. Subtracting one image from the other preserves spatial information that lies between the range of frequencies that are preserved in the two blurred images. Thus, the difference of Gaussians is a band-pass filter that discards all but a handful of spatial frequencies that are present in the original gray scale image.

As a feature enhancement algorithm, the difference of Gaussians can be utilized to increase the visibility of edges and other detail present in a digital image. A wide variety of alternative edge sharpening filters operate by enhancing high frequency detail, but because random noise also has a high spatial frequency, many of these sharpening filters tend to enhance noise, which can be an undesirable arti-fact. The difference of Gaussians algorithm removes high frequency detail that often includes random noise, rendering this approach one of the most suitable for processing images with a high degree of noise. A major drawback to application of the algorithm is an inherent reduction in overall image contrast produced by the operation.

3.4 MFCC

Under clean conditions for audio processing the MFCC's are very useful features. However, performance using MFCC features deteriorates in the presence of noise. There has been an increased effort in recent times to find new features that are more noise robust compared to MFCCs. Features such as, spectro-temporal modulation features are more robust to noise but are computationally expensive. Skowronski and Harris suggested modification of MFCC that uses the known relationship between center frequency and critical bandwidth. They also studied the effects of wider filter bandwidth on noise robustness. Herein, we suggest different modifications to MFCCs that make it more robust to noise

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

without adding prohibitive computational costs. MFCC features approximate the frequency decomposition along the basilar membrane by a short-time Fourier Transform. The auditory critical bands are modeled using triangular filters, compression is expressed as a log function and a discrete cosine transform (DCT) is used to de-correlate the features.

3.5 KNN Classifier

Among several methods for pattern recognition, KNN (k Nearest Neighbors) classifier is one of the most commonly used method and has been applied in a variety of cases. KNN Classifier works as follows. First for each one of the training set elements a classification of it is performed based on various neighborhoods. The k value that maximizes the DC of each classification is found, therefore for each training set there corresponds a particular k value which is considered the best available. Afterwards, for each unknown element, the nearest neighbour is found and its k value is assumed (based on the "optimum" k array). Then, the KNN classifier is applied on that test element, using that k value.
The KNN algorithm is given below

1. Choose k number of samples from the training set to generate initial population (p1).
2. Calculate the distance between training samples in each chromosome and testing samples, as fitness value.
3. Choose the chromosome with highest fitness value store it as global maximum (Gmax).
   a. For i = 1 to L do
      i. Perform reproduction
      ii. Apply the crossover operator.
      iii. Perform mutation and get the new population.
      iv. Calculate the local maximum (Lmax).
      v. If Gmax < Lmax then a. Gmax = Lmax;

Output – the chromosome which obtains Gmax has the optimum K-neighbors and the corresponding labels are the classification results.

IV. EXPERIMENTAL RESULT

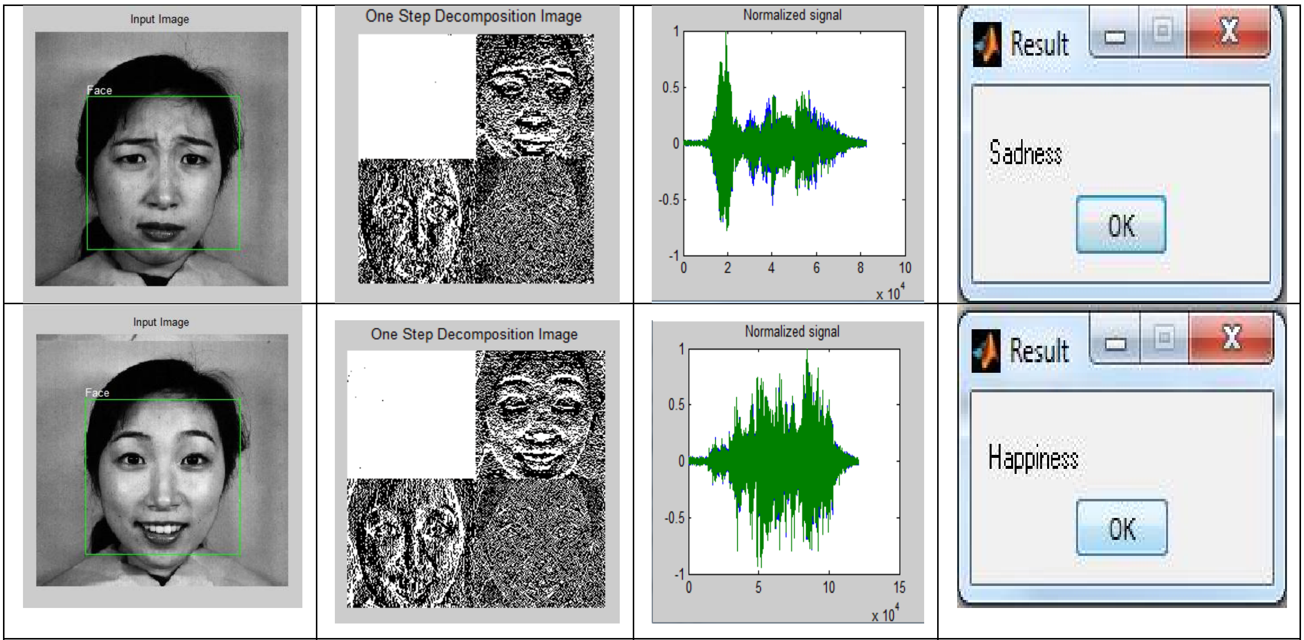
Below figure shows the overall experimental results of the proposed system. We considered Fig.2(a) as input image apply signal processing techniques to convert into time domain signals, after conversion signals are decomposed and considered one decomposition image shown in Fig.2(b). To foreshorten face and emotions accurately and efficiently, will use the MFCC along with LBP (Local Binary Pattern) and DOG for feature extraction, from this will get Normalized signals shown in Fig.2(c), On the basis of result from feature extraction the KNN classifier classifies the activity of our input signal, we obtain output according to the normalized signals as shown in Fig.2(d). Figure 3 represents the overall comparison between existing and proposed system.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016



(a):Input image	(b):Decomposition Image	(c):Normalized Signal	(d):Output

Fig.2: Results of proposed Work

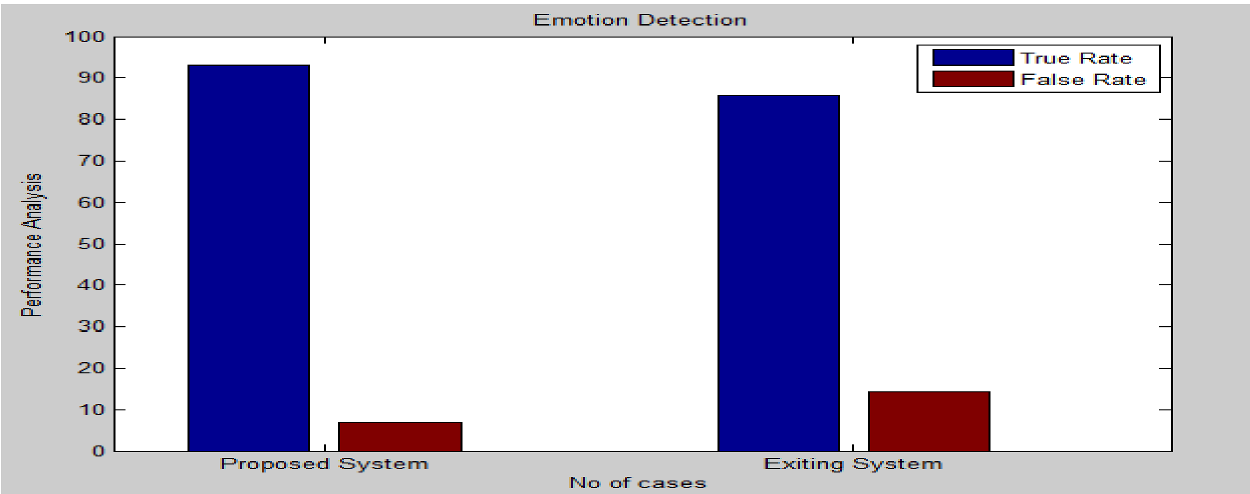


Figure 3: Comparison between proposed and existing system.

V. CONCLUSION

In this project, we introduced and implemented a face and emotion detection algorithm, based on LBP and DOG features. Although LBP feature is simpler, my implementation shows that it is enough to discriminate faces and non faces faster. A Face recognition system can be either verification or an identification system depending on the context of an application.

To foreshorten face and emotions accurately and efficiently, will use the MFCC along with LBP(Local Binary Pattern) and DOG for feature extraction. On the basis of result from feature extraction the KNN classifier classifies the

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

activity of our input signal. Experimental results demonstrate that the proposed approach is superior to other approaches to recognize.

REFERENCES

[1] Viola and Jones, "Face Detection Practically Feasible in Real World Applications Such as Digital Cameras and Photo Organization Software", IEEE , pp.1315–1326, 2010
[2] Ekman and Friesen, "Flourishing Facial Action Coding System", IEEE Trans, Vol. 15, issue 6, pp.1555–1561, 2014.
[3] Murthy and Jadon, "Enhanced Method to Recognize the Expression of the Front View of the Face", IJCSMC, Vol. 4, Issue. 7, pg.334 – 340, 2015,
[4] Fasel and Luettin, "Sign-Based Approach Uses a FACS System", Vol. 19, pp. 823 854, 2014.
[5] K. Reddy and M. Shah, "Recognizing 50 Human Action Categories of Web Videos," Machine Vis. Applications, vol. 24, no. 5, pp. 971–981, 2013.
[6] A. Prest, C. Schmid, and V. Ferrari, "Weakly supervised learning of interactions between humans and objects," IEEE Trans. Pattern Anal. Mach. Intel, vol. 34, no. 3, pp. 601–614, 2012.
[7] J. Sun, Y. Mu, S. Yan, and L.-F. Cheong, "Activity Recognition Using Dense Long-Duration Trajectories," in ICME'10, 2010, pp. 322–327.
[8] C. P. Papageorgiou, M. Oren, and T. Poggio, "A General Framework for Object Detection.", Proc. of ICCV, pages 555-562, 1998.
[9] H. A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection.", IEEE Trans, PAMI, 20(1):23-38, 1998.