International Journal of Scientific & Engineering Research, Volume 4, Issue 11, November-2013 174
ISSN 2229-5518

Design and Implementation of Real Time Facial Emotion Recognition System

Nishant Hariprasad Pandey, Dr. Preeti Bajaj, Shubhangi Giripunje

Abstract— Computer application such as facial and emotion recognition system used for automatically identifying or verifying a person from a digital image or a video frame from a video source. Facial emotions are naturally used by human beings as one way to share their opinions, cognitive, emotions, and intentions states with others and are therefore important in natural communication. One of the ways to achieve this is by Feature based recognition, which is based on the extraction of the properties of individual organs located on a face such as eyes, mouth and nose, as well as their relationships with each other. Haar cascade are used for detection of eyes, mouth and face. Principal component analysis is used for determining different emotions like sad, happy, surprise, neutral in OpenCV.

Index Terms— Communication, Facial emotions, feature extraction, Haar cascade, OpenCV , Principal component analysis ,Real time.

1 Introduction

Facial image processing is one of the most important ones in Computer vision research topics. It consist challenging areas like face detection, face tracking, pose estimation, and face recognition, facial animation and expression recognition. It is also required for intelligent vision-based human computer interaction (HCI) and other applications. The face is our primary focus of attention in social intercourse; this can be observed in interaction among animals as well as between humans and animals (and even between humans and robots [1]).Automatic face detection and recognition are two challenging problems in the domain of image processing and computer graphics that have yet to be perfected. Manual recognition is a very complicated task where it is vital to pay attention to primary components like: face configuration, orientation, location where the face is set (relative to the body), and movement (i.e. traces out a trajectory in space). It is more complicated to perform detection in real time. Dealing with real time capturing from a camera device, fast image processing would be needed [2].The ability to automatically recognize the state of emotion in a human being could be extremely useful in a wide range of gelds. A few examples could be the monitoring of workers performing critical task, an interacting video-game or an active safety device for a vehicle driver. The latter application could be really useful to alert a driver about the onset of an emotional state that might endanger his safety or the safety of others, e.g., when his degrees of attention decay below a given threshold [3]. The six universal emotions across different human cultures comprises of happiness, sadness, fear, disgust, surprise and anger.

Facial Emotion recognition can be done using various algorithms such as using Support Vector Machine, Gabor Features and Genetic Algorithm [4], Mean Shift Algorithm [5], etc. Organization of paper is such that Section II gives the design flow of the system. Section 3 gives description of face detection and algorithm used for this purpose. Section 4 explains the extraction of various facial features. Section 5 is related with facial emotion recognition with the help of principal component analysis. Section 6 comprises of results of different feature extractions. Section 7 concludes the paper with future aspect related with facial emotion recognition. Section 8 comprises of references.

2 Design flow of the system

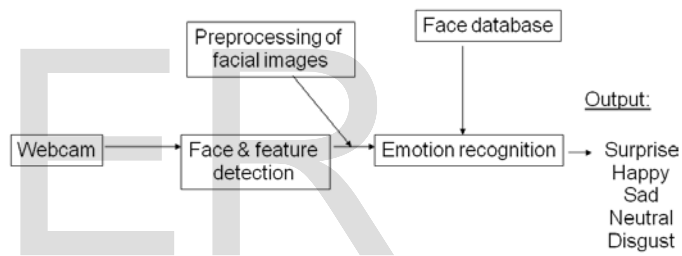

Figure 1: Block Diagram of the system.

Webcam is used for capturing input image frame by frame as required. After capturing input image, detection of face and feature using Haar cascade classifiers is done. Preprocessing (eg:histogram equalisation,contour detection etc) of facial images is necessary for maintaining the efficiency of the system.Atlast emotion of person is identified with the help of Principal component analysis.

3 Design flow of the system

Face detection is the general case of face localization in which the locations and sizes of a known number of faces (usually one) are interpreted. Earlier work is mainly with upright frontal faces, several systems have been developed that are able to detect faces fairly accurately with in-plane or out-of-plane rotations in real time. Although a face detection module is typically designed to deal with single images, its performance can be further improved if video stream is available [6].

Haar-like features developed using Haar wavelet is having edge over other with respect to its calculation speed was introduced by Viola and Jones [7]. Haar-like feature uses contiguous rectangular regions at a particular location in a detection

ISSN 2229-5518

window, adds up the pixel intensities in each region and calculates the difference between these sums. This difference helps to categorize subsections of an image. The area of the eyes is darker than the area of the cheeks among all faces. Therefore a common Haar feature for face detection is a set of two contiguous rectangles that present above the eye and the cheek area. The location of these rectangles is defined relative to a detection window that acts like a bounding box to the face. During detection a window is moved over the input image and for each subsection of the image the Haar-like feature is then calculated. This difference is then compared to a learned threshold that separates non-face from face [7].

The Haar cascade classifier described below has been initially proposed by Paul Viola *Viola01* and improved by Rainer Lienhart *Lienhart02*. First, a classifier (namely a cascade of boosted classifiers working with Haar-like features) is trained with a few hundred sample views of a particular object (i.e., a face or a car), called positive examples, that are scaled to the same size (say, 20x20), and negative examples - arbitrary images of the same size.

After a classifier is trained, it can be applied to a region of interest (of the same size as used during the training) in an input image. The classifiers output a “1” if the region is likely to show the object (i.e., face/car), and “0” otherwise. To search for the object in the whole image one can move the search window across the image and check every location using the classifier. The classifier is designed so that it can be easily “resized” in order to be able to find the objects of interest at different sizes, which is more efficient than resizing the image itself. So, to find an object of an unknown size in the image the scan procedure should be done several times at different scales.

The word “cascade” in the classifier name means that the resultant classifier consists of several simpler classifiers (stages) that are applied subsequently to a region of interest until at some stage the candidate is rejected or all the stages are passed. The word “boosted” means that the classifiers at every stage of the cascade are complex themselves and they are built out of basic classifiers using one of four different boosting techniques (weighted voting). Currently Discrete Adaboost, Real Adaboost, Gentle Adaboost and Logitboost are supported. The basic classifiers are decision-tree classifiers with at least 2 leaves. Haar-like features are the input to the basic classifers, and are calculated as described below. The current algorithm uses the following Haar-like features:

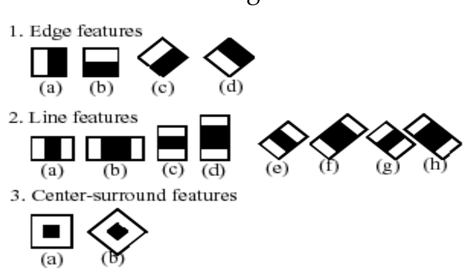

Figure 2: Haar Features.

The feature used in a particular classifier is specified by its shape (1a, 2b etc.), position within the region of interest and the scale (this scale is not the same as the scale used at the detection stage, though these two scales are multiplied). For example, in the case of the third line feature (2c) the response is calculated as the difference between the sum of image pixels under the rectangle covering the whole feature (including the two white stripes and the black stripe in the middle) and the sum of the image pixels under the black stripe multiplied by 3 in order to compensate for the differences in the size of areas. The sums of pixel values over rectangular regions are calculated rapidly using integral images [8]. The integral image is an array containing the sums of the pixels’ intensity values located directly to the left of a pixel and directly above the pixel at location (x, y) inclusive [9].

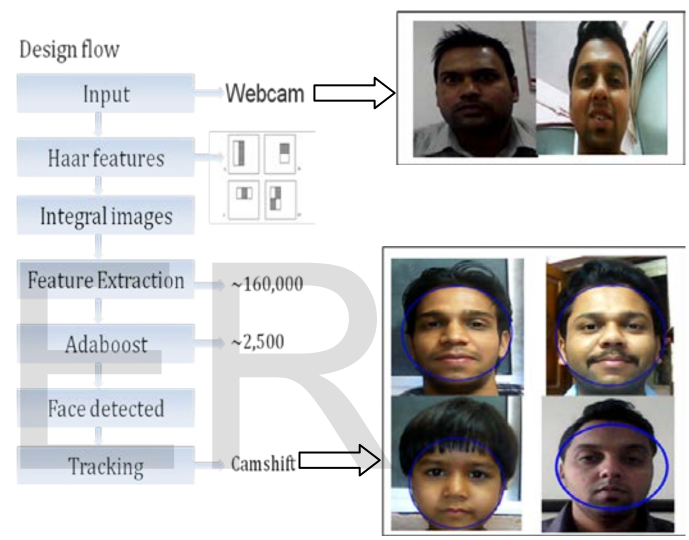

Figure 3: Design flow of Face detection

4 Facial Feature Detection

Detecting human facial features, such as the mouth, eyes, and nose require that Haar classifier cascades first be trained. In order to train the classifiers, this gentle AdaBoost algorithm and Haar feature algorithms must be implemented. Fortunately, Intel developed an open source library devoted to easing the implementation of computer vision related programs called Open Computer Vision Library (OpenCV). The OpenCV library is designed to be used in conjunction with applications that pertain to the field of HCI, robotics, biometrics, image processing, and other areas where visualization is important and includes an implementation of Haar classifier detection and training [10].

OpenCV [11] make available open source libraries which can be used for different real time image processing applications. Object detector in the OpenCV [11] which is based on Haar-like features was used to detect different features such as nose, eyes, mouth, full body, frontal face, etc.

Haarcascade_frontalface_alt.xml for face detection, Haarcas-

ISSN 2229-5518

cade_eye_tree_eyeglasses.xml for eye detection, Haarcascade_mcs_mouth.xml for mouth detection is used for feature detection.

4.1 Eye Detection

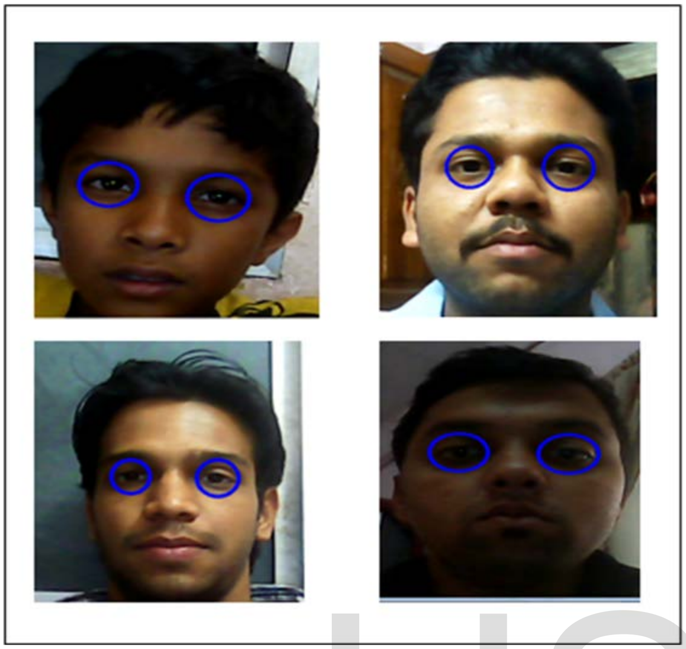

Figure 4: Example of Eye detection

4.2 Mouth Detection

Figure 5: Example of a Mouth detection

4.3 Face,Eye and Mouth Detection

Figure 6: Example of Face, Eye and Mouth detection

However the accuracy of detection in real time is slightly less in case of mouth as compared to eyes and face.Each detection of features taken in different environment and illumination condition. Accuracy of detection of features will depend on the position of face with respect to camera.

5 Facial Emotion Recognition

Facial expression analysis can be performed in different Degrees of granularity. It can be either done by directly classifying the prototypic expressions (e.g. anger, fear, joy, surprise…) from face images or, with finer granularity, by detecting facial muscle activities. The latter is commonly described using the facial action coding system (FACS)[12] Facial Emotion Detection means finding the Expression of an image and recognizes which expression it is such as Neutral, Happy, Sad, Angry, Disgust, etc. Figure 5 shows one example of each facial expression. Emotions could be of the positive or negative type. Happiness, surprise, enthusiasm are some of the positive ones while anger, fear and disgust constitute the latter. A neutral emotion is also defined which is neither positive nor negative but possibly signifies the calm state of a system [13].

The technique used for Facial Emotion Detection is Principal Component Analysis, PCA. The Principal Component Analysis, PCA is one of the most successful techniques that have been used to recognize faces in images. PCA has been called one of the most valuable results from applied linear algebra. It is used abundantly in all forms of analysis (from neuroscience to computer graphics) because it is a simple, non- parametric method of extracting relevant information from confusing datasets. With minimal additional effort, PCA provides a roadmap for how to reduce a complex dataset to a lower dimension to reveal the sometimes hidden, simplified structure that often underlie it [14].

Principal Component Analysis (Eigenface) consists of two phases: learning and recognition. In the learning phase, eigenfaces one or more face images for each person's expression which is to be recognized. These images are called the training images. In the recognition phase, eigenface a face image, it responds by telling which training image is "closest" to the new face image. Eigenface uses the training images to "learn" a face model. This face model is created by applying a method called Principal Components Analysis (PCA) to reduce the "dimensionality" of these images. Eigenface defines image dimensionality as the number of pixels in an image. The lower dimensionality representation that eigenface finds during the learning phase is called a subspace. In the recognition phase, it reduces the dimensionality of the input image by "projecting" it onto the subspace it found during learning. "Projecting onto a subspace" means finding the closest point in that subspace. After the unknown emotion image has been projected, eigenface calculates the distance between it and

ISSN 2229-5518

each training image. Its output is a pointer to the closest training image [15].

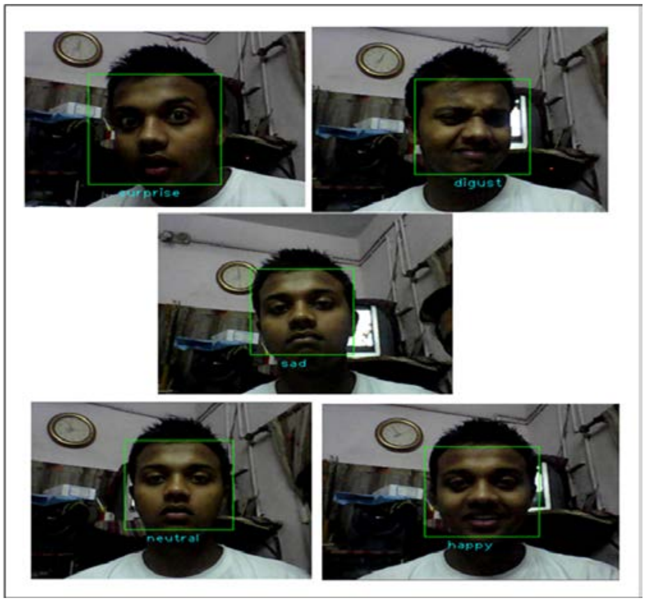

Figure 7: Facial expressions detected using Principal Component Analysis

6 RESULT

Facial extraction method was tested on Ten different people with respect to their face construction and also they differ in age and size. The accuracy shows that the detection of face is most reliable as compared to other features but other features also have good accuracy overall . Detection time of each feature varies with respect to position of face with camera .Detection time will be less if the face is properly align in front of the camera and vice versa.

Table I shows different face feature extraction accuracy and Table II shows different face emotion recognition accuracy.

Table I :Analysis of facial feature detection accuracy

Facial feature detection	*face*	*eyes*	*mouth*	*Face + eyes + mouth*
No of Frames	20	20	20	20
Correct detected frames	19	18	16	14
Accuracy	95%	90%	80%	70%

Table II : Analysis of facial feature detection accuracy

Emotion	*happy*	*sad*	*surprise*	*disgust*	*neutral*
No of Frames	16	16	16	16	16
Correct detected frames	13	12	14	13	15
Accuracy	81%	75%	87%	81%	94%

7 CONCLUSION

In this research work, researchers have discussed facial and emotion recognition system from a variety of aspects such as analytical treatment, features and technologies. Although the concepts are different the implementation was widely done using OpenCV platform. Comparison of several results obtained during facial feature extraction also discussed. The results obtained are promising and have the potential for extending their use in variety of image processing applications such as interactive systems like intelligent transport system, stress recognition etc. The proposed PCA method has the greater accuracy with consistency. The rate of emotion recognition was greater even with the small number of training images it is still fast, relatively simple, and works well in a constrained environment.

REFERENCES

[1] C. Breazeal, Designing Sociable Robots, MIT Press, 2002.
[2] E.M.Bouhabba, A.A.Shafie , R.Akmeliawati ,"Support Vector Machine For Face Emotion Detection On Real Time Basis," 4th International Conference on Mechatronics (ICOM), 17-19 May 2011, Kuala Lumpur, Malaysia ©2011 IEEE.
[3] M. Paschero , G. Del Vescovo, L. Benucci, A. Rizzi, M. Santello, G.Fabbri, F. M. Frattale Mascioli, "A Real Time Classifier for Emotion and Stress Recognition in a Vehicle Driver." IEEE International Symposium on Industrial Electronics (ISIE) ©2012 IEEE.
[4] Seyedehsamaneh Shojaeilangari, Yau Wei Yun, Teoh Eam Khwang, "Person Independent Facial Expression Analysis using Gabor Features and Genetic Algorithm'" 8th International Conference onInformation, Communications and Signal Processing (ICICS) ©2011IEEE.
[5] Peng Zhao-Yi, Wen Zhi-Qiang, Zhou Yu , "Application of Mean Shift Algorithm in Real-time Facial Expression Recognition". International Symposium on Computer Network and Multimedia Technology ©2009 IEEEC. Darwin, The Expression of the Emotions in Man and Animals, Oxford University Press, Third edition, 2002.
[6] I. Cohen, F. Cozman, N. Sebe, M. Cirello, and T.S. Huang.Semi- supervised learning of classifiers: Theory, algorithms, and their applications to human-computer interaction. IEEE Trans. On Pattern Analysis and Machine Intelli-

gence, 26(12):1553–1567, 2004.
[7] Viola and Jones, "Rapid object detection using a boosted cascade of simple features", Accepted Conference On Computer Vision and Pattern Recognition, 2001.
[8] "object detection",opencv.willowgarage.com.
[9] Phillip Ian Wilson , Dr. John Fernandez, "Facial Feature Detection Using Haar Classifiers", Journal of Computing Sciences in Colleges, 2006.
[10] Open Computer Vision Library Reference Manual. Intel Corporation, USA, 2001.
[11] Intel, "Open computer vision library," www.intel.com/technology/computing/opencv/,2008.
[12] P. Ekman and W. V. Friesen. Facial Action Coding System:A Technique for the Measurement of Facial Movement. ConsultingPsychologists Press, Palo Alto, California, 1978.
[13] Dong-Hwa Kim, Peter Baranyi, Nair, "Emotion dynamic express by fuzzy function for emotion robot,"2nd International Conference on Cognitive Infocommunications (CogInfoCom), 2011,
[14] J.shlens,"A Tutorial on Principle Component Analysis," reported by systems Neurobiology Laboratory, Salk Institute for biological studies La Jolla, CA 92037 and Institute for Nonlinear Science, University of California, San Diego L Jolla, CA 92093-0402, December 10,2005.
[15] Article on Face Recognition With Eigenface from Servo Magazine , April 2007.
[16] Expression recognition for intelligent tutoring systems," IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08.
[17] Nilesh U. Powar, Jacob D. Foytik, Vijayan K. Asari, Himanshu Vajaria, "Facial Expression Analysis using 2D and 3D Features," Aerospace and Electronics Conference (NAECON), Proceedings of the 2011 IEEE National ,pp. 73 - 78.
[18] B. Fasel, J. Luettin, "Automatic facial expression analysis: a survey", Pattern Recognition, vol. 36, no. 1, pp. 259-275, January 2003.
[19] M.Turk, W. Zhao, R.chellappa, "Eigenfaces and Beyond", Academic press, 2005.