

Implementation of an Image Processing Algorithm on a Mobile Robot

#¹P.S.Lende, #²Dr.Y.P.Reddy

¹pratiksha12345@gmail.com

²ypreddy.scoe@sinhgad.edu



#¹ Appearing in ME-II year 2014/15 (Mechatronics), Sinhgad College of Engineering, Pune Maharashtra, India

#² Professor, Department of Mechanical Engineering, Sinhgad College of Engineering, Pune Maharashtra, India

ABSTRACT

Facial expressions are very important issue in computer vision, and widely used in the behavioral interpretation of emotions, cognitive science, and social interactions. In this paper, an Image Processing Algorithm is proposed to implement on a Mobile Robot for its use in constrained environment. These mobile robots can be used in security to detect human faces and their expressions. For this, a novel method in a real time approach to emotion recognition through facial expression is presented in live video frame. An automatic facial feature tracker to perform face localization and feature extraction is applied. The facial features in the images are used as input to a Support Vector Machine Classifier. The source of this application will be OpenCV library. For facial expression recognition, initially face is detected, then the facial features are extracted and the facial expressions (emotions) are recognized using Support Vector Machines (SVM). A novel multiclass SVM system of classifiers is used for facial emotion detection and to recognize "Happy" emotion from other emotions. This paper presents the application of the machine learning system of support vector machines to the recognition and classification of facial expressions in both, stored images as well as live captured frame. The robotic system used for experimental validation of the Image Processing Algorithm includes a web camera mounted on a mobile robot.

Keywords— Face detection, OpenCV, Support Vector Machines (SVM), facial expression recognition.

ARTICLE INFO

Article History

Received : 18th November 2015

Received in revised form :
19th November 2015

Accepted : 21st November , 2015

Published online :

22nd November 2015

I. INTRODUCTION

Face is the primary focus of attention in social intercourse, playing a major role in conveying identity and emotion. The detection of faces and the interpretation of facial expression under varying conditions is an everyday task for humans, which is fulfilled without effort. The identity, age, gender as well as the emotional state can be seen from some ones face. The impression which was taken from a displayed expression will affect interpretation of the spoken word and even attitude towards the speaker himself. Humour and sympathy are just two examples for essential information that are primarily communicated via facial expressions.

Facial expression plays an important role in the cognition of human expressions. Recognition of facial expressions could be a vital component of natural human-machine interfaces. It might also be used in behavioral affective state, cognitive science, social interaction and clinical practice. Many experiments has shown that facial expressions played a very important part in the process of verbal communication between speakers. Mehrabian [17] indicated that the 7% of human communication information was communicated by linguistic language (verbal part), 38% by paralanguage (vocal part) and 55% by facial expression. Face detection can be performed based on several clues such as skin color, motion, facial/head shape, facial appearance or a

combination of these parameters. A system has been developed by the authors using non

intrusive approach making use of simple webcam, digital camera and video camera which can be mounted on staff cabin, class rooms, prime locations, crowded malls, roads etc. to track a human face. While tracking a human face, the camera extracts facial features like eye and lips. By analyzing the state of these features, Facial action/gestures and Emotions are recognized. Ekman and Friesen [4] had defined six basic emotions (happiness, sadness, fear, disgust, and anger). Each of these six basic emotions corresponds to a unique facial expression.

Support Vector Machines (SVMs) are a popular machine learning method for classification, regression, and other learning tasks. Author have developed an algorithm using Support Vector Machine that is successfully applied, initially to detect "Happy" emotion. This paper first briefly introduce the SVM algorithm and to extend SVM to process multi-class problem is used, then describe how to get the training and test data from the face database, at last the experiment results and the conclusion is summarized.

II. METHODOLOGY

This gives the brief idea about implementation of an Image Processing algorithm on a Mobile Robot which is available in Mechatronics laboratory in Mechanical Department, SCOE. An algorithm is implemented for Facial Expression Recognition. Facial Expression such as happy, sad, disgust, anger, neutral etc. In this only one expression i.e. happy (smiling) is detected on a Mobile Robot using Support Vector Machine (SVMs). This system is broadly categorized into three stages:

- Face location determination stage
- Feature extraction stage and
- Emotion classification stage

Emotion recognition system will be more robust if emotion together with facial/gesture is integrated to make simultaneous decision as presented in below fig. 1.

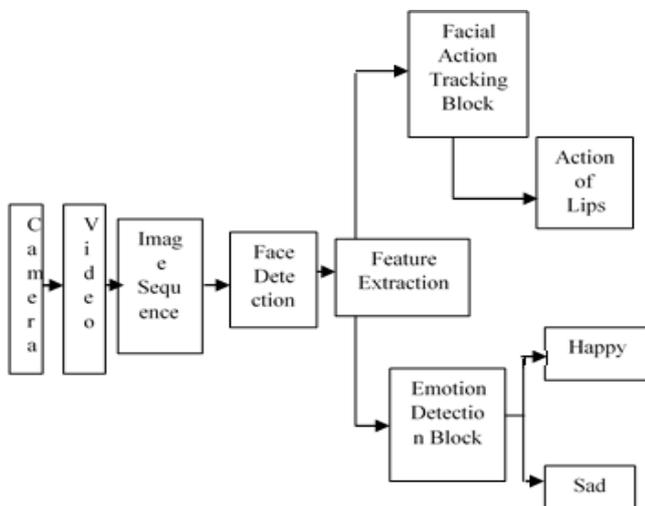


Fig. 1 Block Diagram for Facial Action Tracking and Emotion Recognition From the above block diagram it is clear that, in this system for real time images, firstly through camera images are taken from video frame, initially face is detected, then facial features are extracted and the facial expressions (emotions) are recognized.

A. Face Detection

Detecting the presence of a human face is a complex task in an image due to the possible variations of face. Different sizes, angles and poses a human face might have within the image can cause this variation. The emotions which are detected from human face and different imaging conditions such as illumination and occlusions also affect facial appearances. Face localization and detection is the first step in this work. In this, face is detected through Haar Cascade Classifier. HaarCascade frontal face is the function used in OpenCV library to detect human face from images as well as from live captured frames as shown in below fig. 2.

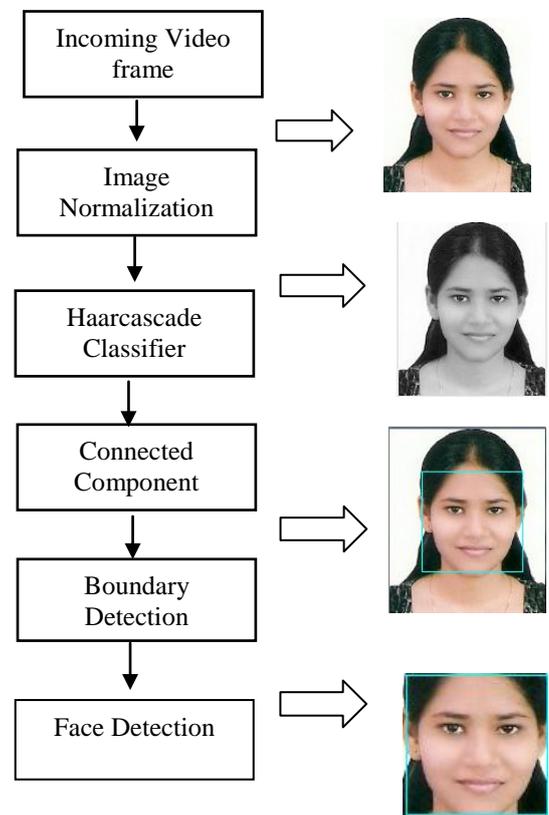


Fig. 2 Flowchart of face detection using Haar classifier

B. Facial Feature Extraction

Facial expressions are the most important information for emotions perception in face to face communication. There are many approaches available for facial feature extraction such as luminance, chrominance, facial geometry and symmetry based approaches, template based approaches, Principal Component Analysis (PCA) and Support Vector Machine (SVMs) based approaches. But in this, Facial features are extracted using Haar Cascade classifier function. There are functions such as Haarcascade mouth, haarcascade lips etc. in OpenCV library to extract mouth,

eye respectively. For extraction of lips haarcascade mouth function is used as shown in Fig. 3.

Facial action is the major source of information for understanding emotional state and intention. The facial actions are recognized in three levels. In the bottom level, a detailed face shape is captured by tracking the facial feature points, which are prominent landmarks such as some meaningful facial behaviours, e.g. mouth open, close mouth etc. surrounding the facial components. Based on the psychological studies of Ekman’s facial action coding system (FACS) [17], authors have used the facial action units (AUs) to characterize these facial behaviours. In the top level, facial expression analysis attempts to recognize facial expressions, i.e., happy and sad. To detect and track changes of facial components in face images, multistate models are developed to extract the geometric facial features. A two-state lip model describes the lip state: open and closed. Given an image sequence, the region of the face and approximate location of individual face features are detected automatically in the initial frame. After the initialization, all face feature changes are automatically detected and tracked in the image sequence.

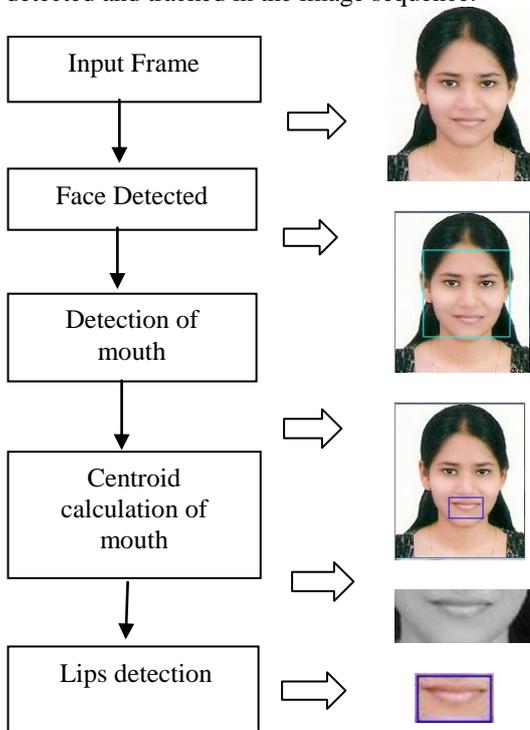


Fig. 3 Flowchart for detection of eyes and lips

C. Emotion Recognition

To analyze the emotions through images and video processing techniques, face location determination and feature extraction was done from images, and then those extracted features were used as input for the classification system which in turn selects a predefined emotion category. Fig. 4 shows how face location determination and feature extraction was done from images. Then these extracted features were used as input for the classification system which in turn selects a predefined emotion category. After the face detection, the colored image is converted into equivalent grayscale image. So that the intensity value of

each pixel is extracted for further feature extraction. As image intensity is at its maximum value, the extracted feature values had fewer errors. Smoothing helps to reduce the noise and also prepare images for further processing. Smoothing can be defined by a convolution operation. The output pixels are obtained by multiplying each neighborhood pixel with a corresponding element of a window and then summing up those products. The Kernel function has been used for smoothing of the images.

Another crucial step in emotion recognition is the design of both a reliable and fast facial features extraction algorithm, whose goal is both - the localization of different facial features (mouth, eyes, eyebrows, nose, etc.) and the analysis of emotional information that they contain (through the shape, the displacement). The features used are based on local spatial position or displacement of specific points and Face region. These facial features extraction and face detection is done by Haarcascade classifier. These Haarcascade classifier are the xml files used in OpenCV library. These files has classifier functions such as Haarcascade frontal face, mouth etc. which are used in algorithm for face detection and features extraction.

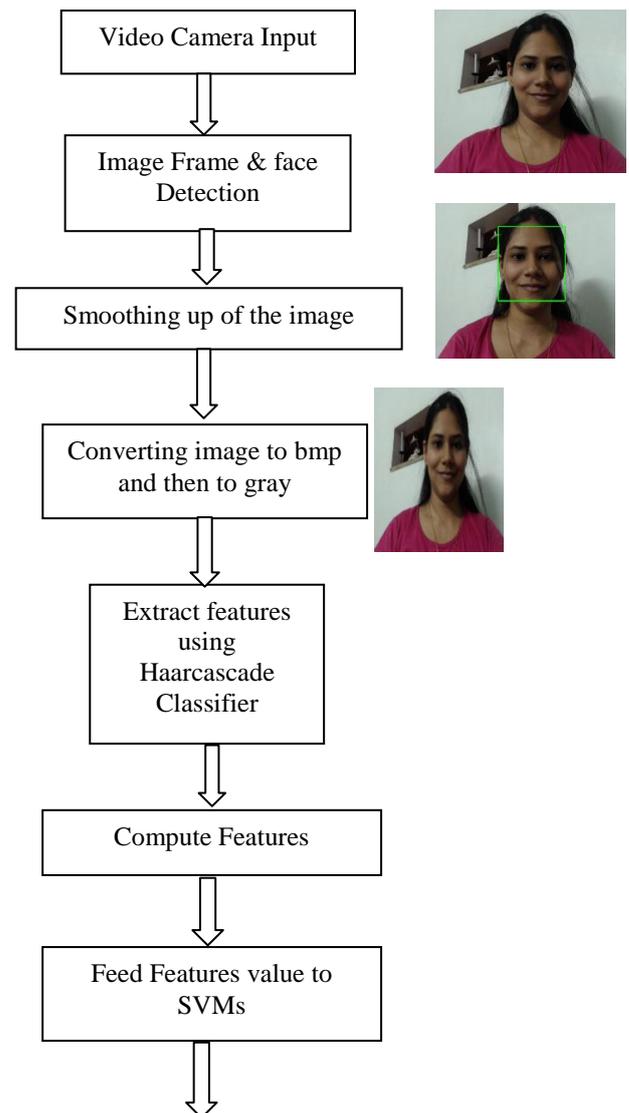




Fig. 4 Flowchart of Emotion Recognition using SVMs

For Emotion Recognition, Support Vector Machine (SVM) System Algorithm is used. Initially image frame are taken as an input through video frame from camera. Then face is detected and features are extracted through Haarcascade Classifier. Then for Facial Expression (emotion) recognition SVMs classifier is used. For SVMs classifier images are initially trained. These trained images creates a database. Number of images are trained in the SVMs classifier and saved as a database for facial expression recognition.

To train the classifier, parameters are set. Parameters are set as a linear for accurate values. When any image is taken as an input through camera, initially face detected than mouth is detected as a feature, then it compares with the trained images in the Smile folder and then the corners of the mouth and the horizontal centre of the mouth are found out and then it resize as a trained size image and then load and save it. It predicts with weight vectors when values are closed to optimal line of smiling image. Thus, smile is detected as a “Happy” emotion. In the comparison when it doesn't get smiling face it is detected as a “Sad” emotion. When no face is detected it compares with the Neutral folder trained images and detect as a “Neutral”.

D. Support Vector Machine Overview

SVMs (Support Vector Machines) are a useful technique for data classification and are still under intensive research [7]. Support Vector Machines classify data through determination of a set of support vectors, through minimization of the Structural Risk. The support vectors are members of the set of training inputs that outline a hyperplane in feature space. This l -dimensional hyperplane, where l is the number of features of the input vectors, defines the boundary between the different classes. The classification task is simply to determinate which side of the hyperplane the testing vectors reside in. Minimizing the structural risk reduces the average error of the inputs and their target vectors. In the description, training data are classified into binary classes. Suppose that training data consists of n samples

$$S = (x_1, y_1), \dots, (x_n, y_n), y_i \in \{+1, -1\}$$

If there is a hyperplane that separates the positive and negative examples, than the point x which lie on the hyperplane satisfy,

$$D(x) = (w \cdot x) + w_0 \quad (1)$$

Where, w is normal to the hyperplane and w_0 is the distance from origin. A separating hyperplane satisfies the constraints that define the separation of the data samples:

$$\begin{aligned} (w \cdot x) + w_0 &\geq +1 & \text{if } y_i = +1 \\ (w \cdot x) + w_0 &\leq -1 & \text{if } y_i = -1, i = 1, \dots \end{aligned} \quad (2)$$

For a given training data set, all possible separating hyperplanes can be represented in the form of equation (2).

Support Vector Machines are based on results from statistical learning theory, pioneered by Vapnik [11], instead of heuristics or analogies with natural learning systems. These results establish that the generalization performance of a learned function on future unseen data depends on the complexity of the class of functions it is chosen from rather than the complexity of the function itself. SVMs perform an implicit embedding of data into a high dimensional feature space, where linear algebra and geometry may be used to separate data that is only separable with nonlinear rules in input space. To do so, the learning algorithm is formulated to make use of kernel functions, allowing efficient computation of inner products directly in feature space, without need for explicit embedding. A nonlinear mapping Φ that embeds input vectors into feature space is given, kernels have the form,

$$K(x, z) = (\Phi(x) \cdot \Phi(z))$$

SVM algorithms separate the training data in feature space by a hyperplane defined by the type of kernel function used. They find the hyperplane of maximal margin, defined as the sum of the distances of the hyperplane from the closest data point of each of the two classes. The SVM algorithm looks for separating hyperplane with the largest margin. The size of the margin bounds the complexity of the hyperplane function and hence determines its generalization performance on unseen data. The SVM methodology learns nonlinear functions of the form:

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + w_0\right)$$

Where the α_i are Lagrange multipliers of a dual optimization problem. It is possible to show that only some of the α_i are non-zero in the optimal solution, namely those arising from training points nearest to the hyperplane, called support vectors. The sign function determines whether the predicted classification $f(x)$ comes out positive or negative. SVMs provides a generic mechanism to fit the surface of the hyperplane to the data through the use of a kernel function. The user may provide a function, such as a line, polynomial, or sigmoid curve, to the SVM, which selects support vectors along the surface of this function. In this, kernel used during the evaluation of SVM-based expression classification is linear function i.e. $(x \cdot z)$.

III. SIMULATION

The simulation of Emotion Recognition is done in Java software with OpenCV library included in it. An application is created in Java for Face detection, Feature extraction and Facial expression recognition. For Smile detection an algorithm is created in Java using SVMs. For this three Application Windows are created. The Application Window as shown in Fig. 5 below.



Fig. 5 Structure of Application Window

A. Application Window

The Structure of Application Window consists of five button for user: Camera, Browse, Capture, Detect and Clear. It has label as “Smile Detection”.

The Camera button is used for live video frame images. The Browse button is used for stored images. Images which are stored are used for smile detection. The Capture button is used when camera is switch on and captures the image. Then Detect button is used to detect face, features and emotion from an image. And the clear button is used to clear the image from the screen.

B. Input Window

Then the Input Window is created when image is used as an input for classification system to recognize emotions. Thus Application window with input image is created. There are two types of images taken as input:

- 1) Stored image
- 2) Image from live video frame

Thus Application window with input image is created as shown in below Fig. 6.



Fig. 6 Application window with input image

C. Output Window

Output window is created when image is taken as input, then that is detected. In detection initially face is detected,

then features are extracted using Haar classifier then emotions are recognized using SVMs classifier. Thus output is given out as “Face detected” or “No Face detected” and emotion recognized as “Happy” ,”Sad” or “Neutral”. When no face is detected it show “Neutral Face”. Thus Application window with results is as shown in the below Fig. 7.



Fig. 7 Application window with results

IV. EXPERIMENTAL RESULTS

The simulation which is done in Java software has to be implemented on a Mobile robot which is available in Mechatronics laboratory in Mechanical Department, SCOE. The Mobile Robot for Emotion recognition is as shown in below Fig. 8.



Fig. 8 Mobile Robot for emotion recognition

This Mobile Robot is a general purpose research robot. It runs on a 24V battery, the power for the robot is managed using the power distribution board which regulates and distributes power to all electronic components of the robot. The robot is controlled through a central computer Atom PC. Atom runs on Microsoft Windows 7. The robot can be fundamentally broken down into three blocks where each block is connected to the Atom-PC. The major blocks of the Chiron are namely

- Power management Unit
- Display Unit
- Motion Control Unit

These blocks further consist of the microcontroller base-boards and sensors. The robot is built by following parts which are briefly described below:

A. Atom PC

It is a central computer which runs on Microsoft Windows 7. The above blocks are connected to Atom PC. Atom-PC is also connected with a webcam which can capture images at high resolution. Some port and peripheral of Atom-PC are left open; thus the user can connect additional component according to the application demand.

B. 24V Battery

The robot is powered by a 24V/7A battery so that it can run longer in a single charge cycle. Though the battery powers the whole robot including the Atom-PC, it is not connected directly. The battery only works as input supply to power distribution board (PDB); which act as power management unit.

C. Power Distribution Board

Power Distribution Boards is intended for supplying power to all robot components (5V → Cheetah-CB, Hydra), (24V → Cheetah, Motor). It is a LPC 1751 based board which has 3.3V, 5V and 12V output channels with max current of 10A.

D. EduARM

EduARM is cortex M3 core LPC 1768 based development board. It has 2.8" touch screen TFT which is being used for display of battery, robot status and informative messages.

E. Hydra

Hydra is main control board for motion control of robot. It is a cortex M3 based LPC1768 based board. Currently, 8 Ultrasonic and 6 IR based sensor are connected to the board. Hydra also sends commands to Cheetah-CB for motor controlling.

F. Cheetah CB

Cheetah-CB is the motor controller board. It accepts the command from Hydra for both the motors and sets the speed for both motors accordingly. It takes feedback from encoder which is connected to motor shaft and achieves desired speed precisely. Cheetah-CB has dedicated connectors for Cheetah for ease of use.

G. Cheetah

Cheetah is the single channel DC motor driver. It can supply maximum 8.6A current to the motor. Cheetah is compatible with Cheetah-CB.

H. Ultrasonic Sensor

Ultrasonic sensor module is used for measuring the distance between two objects. The sensor works on the principle similar to radar or sonar. The sensor produces the high frequency wave and evaluates the echo received back by the sensor. Interface required for sensor reading is GPIO.

I. IR Sensor

Sharp IR Sensor - GP2Y0A21YK0F is a distance measuring unit, can be used for bump detection for robots path. It has measuring range of 10 to 80 cm. The sensor gives analog output voltage proportional to the distance measured. Interface required for sensor reading is ADC.

J. Webcam

A webcam has been integrated with the robot, which can capture images. The webcam mounted on the robot is Logitech C920. The webcam support a high resolution image of 15 MP.

This webcam capture images from video frame. The captured image is then used to detect face, extract facial features and then recognize facial expressions. The HD Pro webcam C920 is shown in the below Fig. 9.



Fig. 9 HD Pro Webcam C920

The syst is been taken at different locations for different people and using HP Pro webcam C920. Each Image has size 640 x 480. The system gives 92.07% Accuracy for Facial Action units using Haar classifier and 87.14% Accuracy for Emotion recognition.

TABLE I
RESULTS OF DIFFERENT FACIAL ACTION UNITS

Facial action units	No. of frames	Correct Detected	Error frames	% Accuracy
Face	42	40	2	95.23
Mouth open	34	31	3	91.17
Mouth closed	25	22	3	88.00
Total	101	93	8	92.07

TABLE III
RESULTS OF DIFFERENT EMOTION RECOGNITION

Expressions	No. of frame	Correct Detection	Error frames	% Accuracy
Happy	40	35	5	87.50
Sad	30	26	4	86.66
Total	70	61	9	87.14

V. CONCLUSION

The Image Processing Algorithm was implemented on a Mobile Robot available in Mechatronics laboratory for Facial Expression Recognition. Thus camera is mounted on the robot and experiment is carried out in Atom PC processor which runs on Microsoft Windows 7. Facial expression recognition was carried out in three stages as: Face detection, Facial Features extraction and Facial expressions (emotion) recognition. For face detection and features extraction Haar classifier is used. And for Emotion recognition we have used Support Vector Machine (SVMs) classifier. SVMs algorithm is implemented for Emotion recognition. For this images are trained and then comparison is made between images. SVMs classify data through determination of a set of support vectors.

Kernel function is used for smoothing of the image. SVMs provides a mechanism to fit the surface of the hyperplane to the data through the use of a kernel function. Linear function is used for classification. Thus, the accuracy is 92.07% for facial action units and for emotion recognition is 87.14%.

ACKNOWLEDGEMENT

It was my pleasure in presenting a paper on dissertation "Implementation of an Image Progressing Algorithm on a Mobile Robot". This dissertation helped in learning many things from technology as well as had practical experience.

I am very glad that I got the opportunity to work under the guidance of Dr. J.L. Minase and very thankful for his methodological support. I am thankful to him for his timely guidance, every help about project and throughout my M.E. course. I am thankful for his support in presenting this paper.

I am also thankful to all my family and friends who have directly or indirectly helped me in this.

REFERENCES

[1] F. Cohn, A. J. Zlochower, J. J. Lien, and T. Kanade, "Automated Facial Expression Recognition Based on FACS Action Units", *Psychophysiology*, 36:35-43, 1999.

[2] T. Kanade, J. Cohn, and Y. Tian., "Comprehensive database for facial expression analysis", in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (FG'00)*, pp. 46-53, 2000.

[3] 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* 22(12): pp. 1424-1445, 2000.

[4] Turk, Matthew A, and Alex P. Pentland, "Face Recognition using eigenfaces", *Computer Vision and Pattern Recognition*, 1991. *Proceedings CVPR'91*, IEEE Computer Society Conference on IEEE, 1991.

[5] Philipp Michel, Rana El Kaliouby, "Real Time Facial Expression Recognition in Video using Support Vector Machines", *ICMI'03*, November 5-7, Vancouver, British Columbia, Canada, 2003.

[6] Hung-Hsu Tsai, Yen-Shou Lai and Yi-Cheng Zhang, "Using SVM to design Facial Expression Recognition for Shape and Texture Features", *International Conference on Machine Learning and Cybernetics*, Qingdao, 11-14 July 2010.

[7] Irene Kotsia and Ioannis, "Facial Expression Recognition in Image Sequences using Geometric Deformation features and Support Vector Machines", *IEEE Transaction on Image Processing*, vol. 16, No. 1, January 2007.

[8] Paul Viola and Michael Jones, "Rapid Object Detection using a Boosted Cascade of Simple features", *IEEE*, 2010.

[9] Praseeda Lakshmi.V., Dr. M.Sasikumar, "Analysis of Facial Expression using Gabor and SVM", *International Journal of Recent Trends in Engineering*, vol.1, No. 2, May 2009.

[10] G.Hemalatha and C.P.Sumathi, "A Study of Techniques for Facial Detection and Expression Classification", *International Journal of Computer Science & Engineering Survey (IJCSES)* vol. 5, No. 2, April 2014.

[11] Guodong Guo, Stan Z. Li, and Kapluk Chan, "Face Recognition by Support Vector Machines", *Proceeding 4th IEEE International Conference*, 2000.

[12] G.U.Kharat and Dr. S.V.Dudul, "Human Emotion Recognition System using Optimally Designed SVM with different Facial Feature Extraction Techniques", *WSEAS Transactions on Computer*, Issue 6, vol. 7. June 2008.

[13] Massimiliano Pontil and Alessandro Verri, "Support Vector Machines for 3D Object Recognition", *IEEE Transaction on Pattern analysis and machine intelligence*, vol. 20, No. 6, June 1998.

[14] Bernd Heisele, Purdy Ho and Tomaso Poggio, "Face Recognition with Support Vector Machines: Global versus Component-based Approach", *Computer vision*, *IEEE International Conference on* vol. 2, 2001.

[15] Deepak Ghimire and Joonwhoan Lee, "Geometric Feature-based Facial Expression Recognition in Image Sequence using Multi-Class AdaBoost and Support Vector Machines", *ISSN*, 2013.

[16] S.P. Khandait, Dr. R.C. Thool, P.D. Khandait, "Automatic Facial Feature Extraction and Expression Recognition based on Neural Network", *International Journal of Advanced Computer Science and Application*, vol. 2. No. 1, January 2011

[17] Aleksic. P.S., Aggelos K. Katsaggelos, "Automatic facial expression recognition using facial animation parameters and multistream HMMS", *IEEE Transactions on Information Forensics and security*, vo. 1. No. 1, March 2006.

