Vol.4. No. 1, March, 2015
95 - 105,ISSN : 0975-5446

Automatic Detection of Facial Expressions from Video Streams

THORAT Bali A
Vision & Intelligent System Lab, Dept of CS and IT,
Dr. B.A.M.U.
Aurangabad (MS) India.
balithorat@gmail.com

MANZA Ganesh R
Planning and Statistics Dept,
Dr. B.A.M.U.
Aurangabad (MS) India.
ganesh.manza@gmail.com

YANNAWAR Pravin L
Vision & Intelligent System Lab, Dept of CS and IT,
Dr. B.A.M.U.
Aurangabad (MS) India.
pravinyannawar@gmail.com

Abstract:

Facial expression is one of the most powerful, natural, and immediate means for human beings to communicate their emotions and intensions. Facial expression analysis refers to computer systems that attempt to automatically analyzed and recognize facial motions and facial feature changes from visual information. In this research work we have analyzed changes in facial expression by recognizing facial muscle action units (AUs) and analyzing their temporal behavior, by detecting AUs from face video streams for Happy and Surprise expressions, it is essential to develop automatic expression recognition system to analyze changes in facial expressions based on permanent facial features such as brows, eyes and mouth. This research work attempts to recognize fine-grained changes in facial expression based on Facial Action Coding System (FACS) and automatic detection action units (AUs) based on threshold; this evaluation is based on geometrical features and threshold distance between standard points for person-independent facial expression recognition.

Keywords: Face Detection, Facial feature point extraction, Facial expression recognition, Feature extraction.

1 Introduction

Computer has entered in every walks of human life and has contributed significantly in making this world digital. The increasing importance of computers in daily life [1], has emerged an area of Human Computer Interaction (HCI) to the researchers where time was spent on building efficient systems that can interact with human and vice-versa. Developing a HCI system which leverage communication based on expression will be established and which could detect (using face detection) and understand human behavior (using facial expression analysis) in order to respond to their needs or requests. This approach would significantly improve performance of HCI systems over time. Automatic facial expression analysis is an important part of such a system. Human faces contain significant information about emotions and the mental state of a person that can be utilized in order to enable non-verbal communication with computers.

Automatic expression analysis systems attempt to recognize a small set of prototypic expressions (i.e. Smile and Surprise). In everyday life, however, such prototypic expressions occur relatively infrequently [2, 3]. Instead, emotion is communicated by changes in one or two discrete facial features, such as lips corners are pulled obliquely in smile. Change in isolated features, especially in the area of the brows or eyelids, is typical of paralinguistic displays; for instance, raising the brows signals greeting. To capture the subtlety of human motion and paralinguistic communication, automated recognition of fine-grained changes in facial expression is needed [4].

The content of the paper is organized in Seven sections, Section I deals with Introduction, Section II deals with related work, Section III deals with Dataset, Section IV deals with methodology adopted, Section V deals with Action units, Section VI provides detailed experimental analysis, Section VII is Result and Section VIII deals with conclusion of work and references.

2 Related Work

As per Meharabian [5], 55% communicative cues can be judge by facial expression; hence recognition of facial expressions became a major modality. For example, Smart Devices like computer/robots can sense/understand the human's intension from their expression then it will helpful to the system to assist them by giving suggestions or proposals as per their needs. Ekman and Friesen [6] represents 6 basic face expressions (emotions), which are Happy, Surprise, Disgust, Sad, Angry, Fear. Until now much research work has been done on detecting the human faces based on templates, neural networks and example-based techniques. However, some of these methods are computationally expensive, dealt with single frontal view and work on gray scale images.

Govindaraju, S. N. Srihari and D. B. Sher [7], 1990, presented a system, which detects faces in photographs of newspapers, but the approximate size and the expected number of faces must be known in advance. R. Chellappa, C. L. Wilson and S. Sirohey [8], 1995, proposed a methodology on Human and machine recognition of faces. J.L.Crowley and J. Coutaz [9], 1997, introduced a method on Vision for man machine interaction. C. Garcia and G. Tziritas [10], 1999, presented a system on face detection using quantized skin color regions merging and wavelet packet analysis. Lee, Dileep M.R and Ajit Danti [11], proposed LC face model for recognition of human facial expressions. Face Expression approach (As per Anastasios et al. [12]) can be divided into three major steps so that the face in an image is known for further processing, facial feature extraction which is the method used to represent the facial expressions and finally classification which is the step that classifies the features extracted in the appropriate expressions.

Jyh-Yeong[13], 2001 proposed a methods and technique of Neural Network + Rough Contour Estimation Routine (RCER) where they describe radial basis function network (RBFN) and a multilayer perception (MLP) network. Jacob [14], 2006, proposed the Haar+Adaboost method which achieved comparable accuracy to the Gabor+SVM method for AUs of the eye and brow regions, but it performed very poorly for AUs of the mouth. Michal [15], 2008 proposed 2D Gabor filter, Multichannel Gabor filtration scheme used for the detection of salient points and the extraction of texture features for image retrieval applications. Ganesh Manza et.al (2008) has addressed use of composite action unit normalization for laughing face.[16] Manal Abdullahv[17], 2012 proposed Principal Component Analysis which is useful where larger database and less computational time. Murthy [18], 2009 proposed PCA + Eigenfaces, Compared with the facial expression recognition method based on the video sequence, the one based on the static image is more difficult due to the lack of temporal information. Neeta [19], 2010, they use the eyebrow and mouth corners as main "anchor" points. The system, based on a local approach, is able to detect partial occlusions.

3 Database

The aim of this paper is to explore possibility of threshold oriented geometric feature descriptors for recognition of facial expressions and to find the more suitable feature descriptor for robust facial expression recognition. For this work, there is need to record various expressions from different subjects and formulate an expression database *'VISExp'*. '*VISExp*' database was collected in laboratory of vision and intelligent system lab of computer science department. Each expression was recorded for the duration of one second and mode of acquisition was *'Neutral-Expression-Neutral'* constraint without head movement. The subject chosen for data collection were from department of drama, and are professional theatre artist. These videos were collected in standard constrained environment with following precautions:1) Distance between camera and subject, 2) Background, 3) Time for acquisition etc,. There were 59 samples collected for the work which includes four different subjects and each subject poses for several neutral, surprised and happy expressions. Sample of database is shown in figure 1 and collection of dataset is shown in Table 1.

Figure 1

Table 1 'VISExp' Dataset

Subject	Expressions					
	Neutral		Happy		Surprise	
	Video Streams	Frames	Video Streams	Frames	Video Streams	Frames
Subject 1	5	130	3	78	4	104
Subject 2	5	130	7	182	7	182
Subject 3	5	130	5	130	4	104
Subject 4	5	130	5	130	4	104
Total	20	520	20	520	19	494

Other expressions such as 'Sad', 'Disgust', 'Fear', 'Anger' were also being considered at the time for constitution of database. Sampling rate for video stream is *26fps*, this has generated total 520 frames containing neutral expression, 520 frames containing happy and 494 frames containing surprise expression form respective video streams.

4 Methodology

The facial feature recognition system accepts the visual input. The video input is captured with the help of standard cameras. Once the input is acquired, it will be pre-processed and saved into 'VISExp' database. The every stream from database was called for feature extraction and further processing when required. Facial expression analysis includes both measurement of facial motion and recognition of expression. The general approach to automatic facial expression analysis consists of three steps: face acquisition, facial feature extraction and representation, and facial expression recognition.

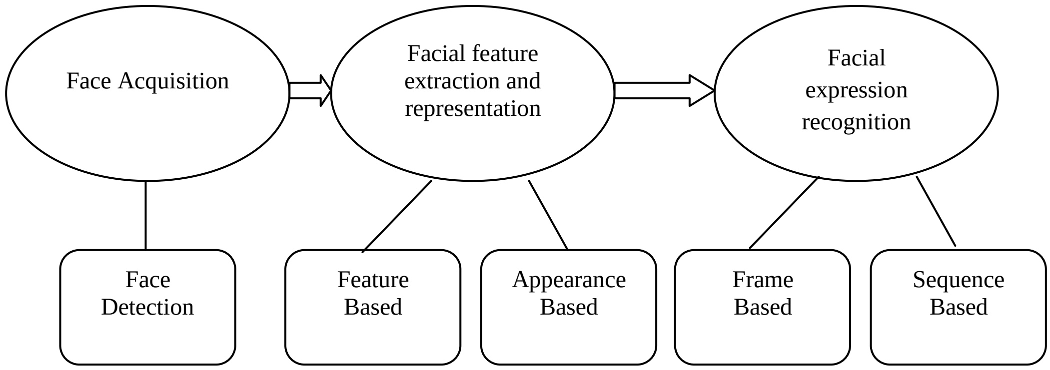

Figure 2 Basic structure of facial expression recognition

4.1 Face Acquisition

Face acquisition is a processing stage and objective is to automatically find the face region for the input video or sequences. It can be through detector to detect face for each frame of the video sequence.

To detect the face image in a scene we make use of a real-time face detection scheme proposed in [20], which represents an adapted version of the original Viola-Jones face detector [21]. The Viola-Jones face detector consists of a cascade of classifiers trained by AdaBoost. Each classifier employs integral image filters, which remind of Haar Basis functions and can be computed very fast at any location and scale. This is essential to the speed of the detector. For each stage in the cascade, a subset of features is chosen using a feature selection procedure based on AdaBoost.

4.2 Facial Feature Extraction and Representation

After the face is located, the next step is to extract and represent the facial changes caused by facial expressions. In facial feature extraction for expression analysis, there are mainly two types of approaches: Geometric feature-based methods and Appearance-based methods. The geometric facial features present the shape and locations of facial components (including mouth, eyes, brows, and nose). The facial components or facial feature points are extracted to form a feature vector that represents the face geometry. With appearance-based methods, image filters, are applied to either the whole-face or specific regions in a face image to extract a feature vector

4.3 Facial Expression Recognition

Facial expression recognition is the last stage. The facial changes can be identified as facial action units or prototypic emotional expressions. In this paper we classified the recognition approaches as frame-based or sequence-based.

5 Action Units

Scientific study of facial expressions began with the team led by Ekman [22]. Ekman and Friesen [23] developed the Facial Action Coding System (FACS) for describing facial expressions. The FACS is a human observer based system designed to describe subtle changes in facial features. FACS consists of 44 action units, including those for head and eye positions. AUs are anatomically related to contraction of specific facial muscles. They can occur either singly or in combinations. AU combinations may be additive, in which case combination does not change the appearance of the constituents, or non-additive, in which case the appearance of the constituent's changes. Automatic recognition of action units is a difficult problem. AUs have no quantitative definitions and as noted can appear in complex combinations.

The AUs are also presented and used in groups as combined complex expression. The groups are based upon the location and/or type of action involved. Upper Face Action units are provided guidelines to code muscular changed appeared in upper face particularly affecting to eyebrows, forehead, and eyelids whereas Lower Face AUs are concerned with five groups that is Up/Down, Horizontal, Oblique, Orbital, and Miscellaneous Actions.

Table 2 Upper Face AUs and some Combinations

AU Number	**FACS Name**
AU 0	Neutral face
AU 1	Inner Brow Raiser
AU 2	Outer Brow Raiser
AU 4	Brow Lowerer
AU 5	Upper Lid Raiser
AU 6	Cheek Raiser
AU 7	Lid Tightener
AU 1+AU 2	Inner & Outer portions of the brows are raised
AU 1+ AU 4	Medial portion of the brows is raised & pulled together
AU 4+ AU 5	Brows lowered and drawn together and upper eyelids are raised
AU 1 + AU 2 + AU 4	Brows are pulled together and upward
AU 1+ AU 2+ AU 5	Brows and upper eyelids are raised
AU 1+ AU 6	Inner portion of brows and cheeks are raised
AU 6+ AU 7	Lower eyelids cheeks are raised
AU 1+ AU 2+ AU 5+ AU 6+ AU 7	Brows, eyelids and cheeks are raised

Table 3 Lower Face AUs and some Combinations

AU Number	**FACS Name**
AU 0	Neutral
AU 9	Nose Wrinkler
AU 10	Upper Lip Raiser
AU 12	Lip Corner Puller
AU 15	Lip Corner Depressor
AU 17	Chin Raiser
AU 25	Lips Part
AU 26	Jaw Drop
AU 27	Mouth Stretch
AU 23+ AU 24	Lips tightened, narrow, and pressed together
AU 9+ AU 17	Nose Wrinkler, Chin Raiser
AU 9+ AU 25	Nose Wrinkler, Lips Part
AU 9+ AU 17 + AU 23+ AU 24	Nose Wrinkler, Chin Raiser, Lips tightened, narrow, and pressed together
AU 10+ AU 17	Upper Lip Raiser , Chin Raiser

AU 10+ AU 25	Upper Lip Raiser , Lips Part
AU 10+ AU 15+ AU 17	Upper Lip Raiser, Lip Corner Depressor, Chin Raiser
AU 12+ AU 25	Lip Corner Puller , Lips Part
AU 12+ AU 26	Lip Corner Puller, Jaw Drop
AU 15+ AU 17	Lip Corner Depressor, Chin Raiser
AU 17+ AU 23+ AU 24	Chin Raiser, Lips tightened, narrow, and pressed together

6 Experimental Work

The block diagram of basic structure of facial expression recognition is as shown in figure 2. The proposed system was designed in light of basic system, which accepts video feed containing expression from VISExp database and passed to pre-processing operation where each video feed was down sampled into discrete 2D image frame with a sampling rate of 26fps. This 2D image vector is further processed for actual extraction of features for expression and thresholds are to be determined so that changes can easily be tracked. The feature extraction process extracts the features of expression by automatically locating face as per figure 3 and face components related to Upper Face (eyebrows, eyelids, forehead) etc as well as Lower facial components (mouth, nose, nostrils) etc. This was done by using 'viola-jones' algorithm.

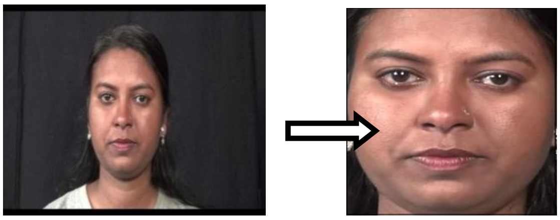

Figure 3: Detection of Face from 2D Frame

The Viola-Jones algorithm was used which was based on Haar-like features comprising of two edge features, line feature and rectangle feature. This also detects eyes, nose and mouth shown in figure 4(a) were based on cascade object detector which was further modified in this research work for adding two more cascaded objects like eyebrows and eyeballs as shown in figure 4(b). These cascaded objects were detected using feature selection procedure based on AdaBoost algorithm. Each detected component from each frames are marked by bounding-box.

(a) (b)

Figure 4 (a) Detected face components, (b) Detected Eyebrows

Once the basic face components have been identified it is necessary to code the distance among these principal components of the face. To do so, the center of each bounding box associated with face components have been earmarked as shown in figure 5 and corresponding pixel oriented distance have been recorded as per figure 6.

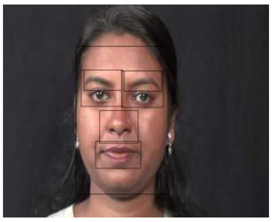

Figure 5: Demarking center of Bounding Box.

The Normalized distance between components are considered as left eye to left eyebrow, right eye to right eyebrow, left and right eye to nose, nose to mouth

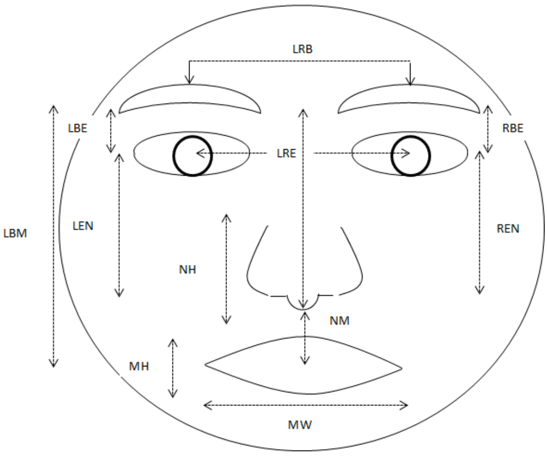

Figure 6: Geometric Face

Where,
LRB – Distance between left to right brow.
LBE – Distance between left brows to left eye.
RBE – Distance between right brows to right eye.
LRE – Distance between left to right eye.
LEN – Distance between left eye to nose.
LBM – Distance between left eyebrow to mouth.
REN – Distance between right eye to nose.
NM – Distance between nose to mouth.
NH – Height of Nose.
MW – Width of mouth.
MH- Height of Mouth.

Figure 6 shows the geometric face which has the normal Distance threshold values as shown in Table 4. The expression is constitutes in change in the muscular moment of face from natural state to expression state. Hence it is important to understand the natural neutral state of face. In order to decide the natural neutral threshold value associated with component distance like {distance between the components of face such as distance between two eyes, distance of nostril from left eyeball & right eyeball, distance between eyeball and eyebrow, distance between nostril and mouth, mouth height & width} the video stream was processed and suitable threshold features were extracted and stored. In this recognition process the threshold value of neutral frames was recorded, then this threshold value was compared with other frames which have different actions i.e. when subject give surprise movement his eyebrows go up i.e the distance between eyes and eyebrows are increases, similarly, when subject is smile his lip corner are pulled obliquely i.e the mouth width are increase.

LRB	LBE	RBE	LRE	LEN	REN	NM	MW
70	13	13	66	79	79	48	80

Table 4 Normal Distance Threshold values for Neutral Expression

Similarly, Table 5 shows the threshold values of surprise movement, when subject is in surprise his eyebrows are raised and sometimes mouth is open.

Table 5 Distance threshold value of Surprise Expression.

LRB	LBE	RBE	LRE	LEN	REN	NM	MW
70	16	16	66	79	79	48	80

The sample frame is as shown in figure 7.

Figure 7 Sample Frame

These distances are plotted as shown in below figure 7i which shows the distance between eyes to eyebrows when subject is in neutral and when subject gives surprise expression.

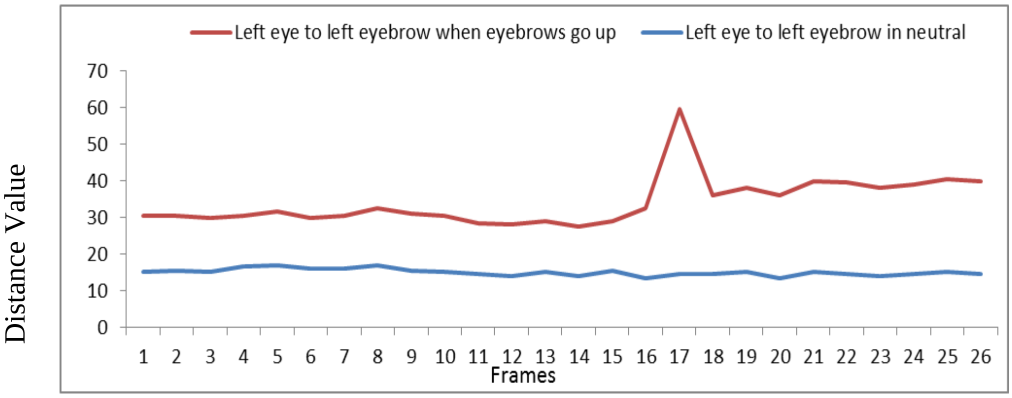

Figure 7i (a)

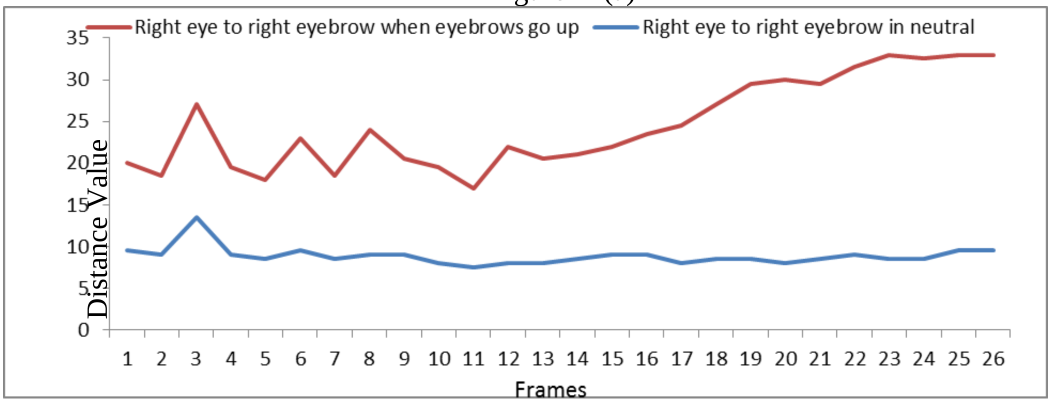

Figure 7i (b)

It shows the variance between neutral frame and active frame. In figure 7i shows the graph of two samples, figure 7i(a) is for neutral frames and active frames i.e when left eyebrows go up and figure 7i(b) is for neutral frames and active frames when right eyebrows go up.
Based on this observation, it recognized that the distance threshold values of eyes and eyebrows are changed and at the same time in some cases the threshold value of mouth is also changed. Hence it recognizes that the Action Unit AU1+AU2+AU5+AU25+AU27 are active, which detect the Surprise expression.

To recognize mouth action unit Sobel edge detection function was used. Using this function the threshold value of neutral face mouth was extracted and compare with smiley frames and detect the Action Unit 12 which refer for Lips corner puller it detects that subject gives smile expression. Figure 8i(a) shows the flow of graph when frames are neutral and Figure 8i(b) shows the flow of graph when frames gives the smile expression.

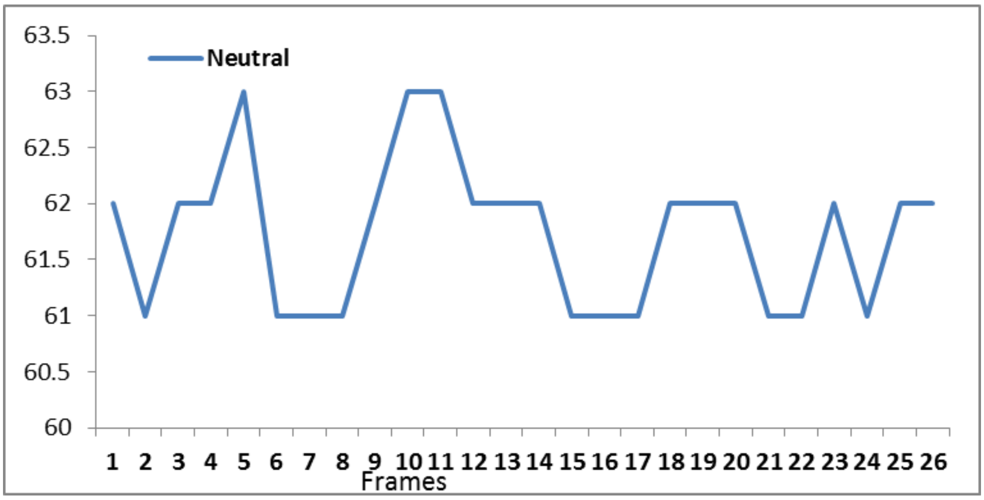

Figure 8i (a)

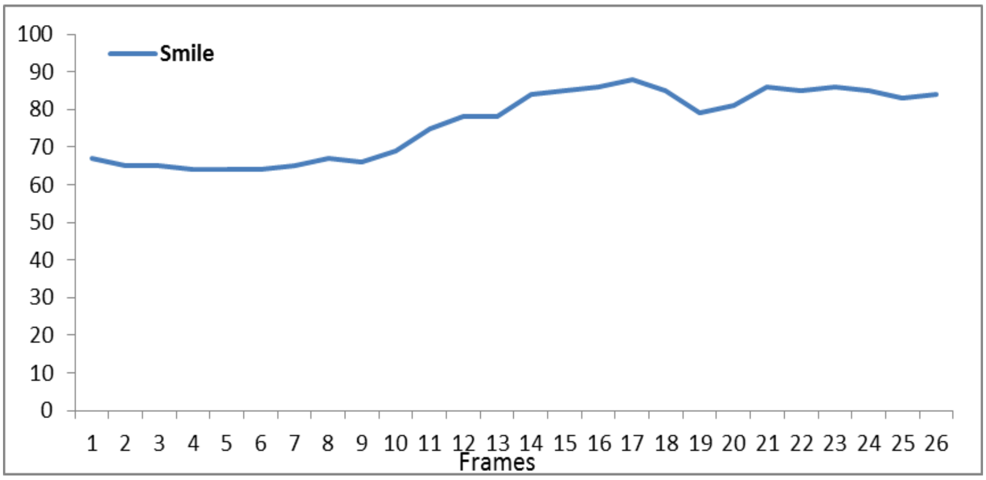

Figure 8i (b)

7 Result

The feature extraction method is tested on face videos of 4 subjects that contain Smile, neutral, and surprise expressions. Here, the results from the Geometric and edge detection method are discussed. Geometric method on eyes and Edge detection on mouth frames using morphological operation and Sobel edge detection function. Thus, the edges have form unique shapes that can be differentiated from one expression to another expressions. Once the parameters were obtained the threshold value was set and fixed the Action Unit for that value and recognizes the facial expression. Table 6 shows the subject wise recognition rate.

Table 6 Over all Recognition Rate for Smile and Surprise Expression

Subject	Recognition Rate %	
	Smile	**Surprise**
Subject 1	96.00%	82.81%
Subject 2	96.00%	88.33%
Subject 3	98.89%	83.72%
Subject 4	82.76%	98.25%
Overall Recognition Rate	93.41%	88.27%

8 Conclusion

In this proposed research work describes the medication of cascade object detector mechanism of ‘Viola-Jones’ for detection of important face component that contributes in formation of facial expression. The result obtained clearly provides confirmation about possibility of using thresholds for detection of expression form a video streams. The prototyped mechanism will be further tested with variety of expressions such as ‘sad’, ‘disgust’, ‘anger’ and ‘fear’.

Acknowledgement

The Authors gratefully acknowledge support by the Department of Science and Technology (DST) for providing financial assistance from Major Research Project sanctioned under *Fast Track Scheme for Young Scientist*, vide sanction number SERB/1766/2013/14 and the authorities of Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS) India, for providing the infrastructure for this research work.

References

[1] B.Fasel and J.Luettin, “Automatic Facial Expression Analysis: A Survey”, Pattern Recognition, vol. 36, no. 2, 2003, pp. 259-275.

[2] K. Mase. Recognition of facial expression from optical flow. IEICE Transc., E. 74(10):3474–3483, 0ctober 1991.

[3] Y. Yacoob and L. S. Davis. Recognizing human facial expression from long image sequences using optical flow. IEEE Trans. On Pattern Analysis and machine Intelligence, 18(6):636–642, June 1996.

[4] Yingli Tian Takeo Kanade Jeffrey F. Cohn, “Recognizing Action Units for Facial Expression Analysis”, December, 1999.

[5] Mehrabian.A, "Communication without Words", Psychology Today,1968. Vo1.2, No.4, pp 53-56.

[6] Ekman, P, Friesen, “Constants across Cultures in the Face and Emotion”, J. Pers. Psycho. WV, 1971, vol. 17, no. 2, pp. 124-129.

[7] C. Garcia and G. Tziritas, Face detection using quantized skin color regions merging and wavelet packet analysis, IEEE Trans. Multimed. 1(3) (1999) 264–277.

[8] R. Chellappa, C. L. Wilson and S. Sirohey, Human and machine recognition of faces: a survey, Proc. IEEE 83(5) (1995) 704–740.

[9] J.L.Crowley and J. Coutaz, Vision for man machine interaction, Robo. Auton. Syst. 19 (1997) 347–358.

[10] Gagliardi C, Frigerio E, Burt DM, Cazzaniqa I, Perrett DI, Borgatti R, Facial expression recognition in Williams syndrome, neuropsychologia.2003;41(6):733-8.

[11] De Silva, Liyanage Chandratilake, Facial emotion recognition using multimodal information, vol 1(1997)pp 397-401.
[12] Anastasios C. Koutlas, Dimitrios I. Fotiadis “A Region Based Methodology for Facial Expression Recognition” page 1.
[13] Jyh-Yeong Chang and Jia-Lin Chen, “Automated Facial Expression Recognition System Using Neural Networks” Journal of the Chinese Institute of Engineers, Vol. 24, No. 3, pp. 345-356 (2001).
[14] Jacob Richard-Whitehill, “Automatic Real-Time Facial Expression Recognition for Signed Language Translation”, Department of Computer Science, University of the Western Cape. May 2006.
[15] Tomasz Andrysiak, Michał Chora´S “Image Retrieval Based On Hierarchical Gabor Filters” Int. J. Appl. Math. Compt. Sci., 2005, Vol. 15, No. 4, 471–480.
[16] Ganesh R Manza, P L Yannawar, “Mathematical Normalization of Laughing face using Action Units”, Proceedings of International Conference on Advances in Computing, Anuradha Engineering College, Chikhali, ISBN:978-81-906457-0-6,Feb 21-22,2008 pp 251-253.
[17] Manal Abdullah1, Majda Wazzan1 and Sahar Bo-saeed “Optimizing Face Recognition Using PCA” International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.2, March 2012.
[18] G. R. S. Murthy, R.S.Jadon “Effectiveness of Eigenspaces for Facial Expressions Recognition” International Journal of Computer Theory and Engineering, Vol. 1, No. 5, December, 2009, pp. 1793-8201.
[19] Neeta Sarode, Prof. Shalini Bhatia, “Facial Expression Recognition”, International Journal on Computer Science and Engineering, Vol. 02, No. 05, 2010, 1552-1557.
[20] I.R. Fasel, B. Fortenberry and J.R. Movellan, “A generative framework for real time object detection and classification”,Int’l J Computer Vision and Image Understanding, vol. 98, no. 1, pp. 181-210, 2005.
[21] P. Viola and M. Jones, “Robust real-time object detection”, Technical Report CRL 20001/01, Cambridge Research Laboratory, 2001.
[22] P. Ekman and W.V. Friesen, "The Facial Action Coding System", San Francisco: Consulting Psychologist Press, 1978.
[23] P. Ekman and W. V. Friesen. The Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press Inc., San Francisco, CA, 1978.