Real Time Implementation of Emotion Detection and Classification system for Autism Patients

Mohana and H. V. Ravish Aradhya

December 19, 2019

Real Time Implementation of Emotion Detection and Classification system for Autism Patients

Mohana
Department of Telecommunication
R.V.College of Engineering
Bangalore, India
mohana.rvce@gmail.com

H. V. Ravish Aradhya
Department of E. C. E
R.V.College of Engineering
Bangalore, India
ravisharadhya@rvce.edu.in

Abstract— **Autism is a neurodevelopment disorder characterized by impaired social interaction with verbal and nonverbal communication. Autism is usually diagnosed in the children of age group 4 to 15 years. 1500 out of 1 million children are diagnosed with autism problems. Many autistic patients have impairment in understanding other people's emotions which hinder their interpersonal communication. There is no effective solution to aid autistic patients in understanding the expression of the people with whom they are communicating. Problem with expression detection system is to be able to make a person independent model and to be able to make an efficient system in classifying the emotions into one of the following expressions such as anger, smile, surprise, neutral, sad etc.**

Mainly two types of approaches to extract facial features are reported geometry-based methods and appearance-based methods. In this paper appearance based method is used for extracting the facial features. This paper describes the different facial expressions of humans, real-time face detection, extraction of features from the face and finally a decision based system is built to classify the different facial expressions. The methodology involves resizing the video in to 256*256 pixels using Lanczos resizing filter, grey scale conversion is carried out by luminosity method, from which face is detected using Viola and Jones algorithm it uses Haar features for face detection, further Eigen values are extracted from the detected face which is used for classification. In this system 215 images have been trained and the range of Eigen value for classification is determined. Hence, from the range of values obtained, efficiency of facial expression detection for smile, anger, surprise, sad and neutral are 90%, 83%, 80%, 80% and 90% respectively. Thus, the autistic patients will be notified with expression detected which will be effective for their communication.

Keywords— Eigen values; luminosity method; Grey scale; Vila Jones; Clustering.

I. INTRODUCTION

Autistic children show less attention to social stimuli, smile and look at others less often, and respond less to their own name. Autistic children differ more strikingly from social norms; for example, they have less eye contact and turn taking, and do not have the ability to use simple movements to express themselves, such as the deficiency to point at things. Three- to five-year-old children with autism are less likely to exhibit social understanding, approach others spontaneously, imitate and respond to emotions, communicate nonverbally, and take turns with others. Older children and adults with autism perform worse on tests of face and emotion recognition. Conventional expression detection system has problems with face localization and feature extraction when the system deals with spontaneous expressions [7]. One of the other problems with conventional expression detection was that neutral faces were overrepresented by an order of magnitude more than any other class. As per the survey, there is no effective system which helps autistic patient in understanding the facial expression. Hence there is a need to build a real time expression detection system. It is evident that for classification some of the standard classification algorithms such as K-means classification, hierarchical classification and density based classification are used, in this paper the emotion classification efficiency is improved by using Eigen values or thresholding algorithm. Faulty rate of the expression detection system is reduced.

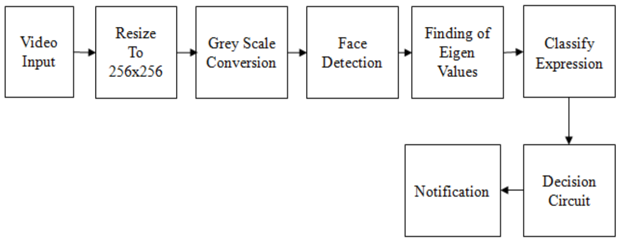

Fig.1. Real time emotion detection and classification system.

Real time emotion detection and classification system as shown in figure 1. It resizes various facial expressions taken from input video in to 256X256 by using lanczos resizing filter. Conversion of RGB video to greyscale using luminosity method, detection of the face using viola and jones algorithm using Haar features, features such as Eigen values are extracted, training of database images and classification is carried out using thresholding algorithm which uses range of Eigen values for classifying 5 different expressions i.e. smile, anger, surprise, neutral and sad. Finally decision system is built to determine the expression and Notifying the patient about the expression.

II. DESIGN AND IMPLEMENTATION

This section describes the implementation details of emotion detection and Classification system

A. Image size reduction

The input image size is reduced to 256x256. The resizing of image is very important since, it helps in reducing the memory required to store the images. The Lanczos filter which is used to smoothly interpolate the value of a digital signal between its samples is the most widely used technique for resizing of images since; it is much faster and less complex than others. Lanczos re sampling or Lanczos filter is a mathematical formula used to smoothly interpolate the value of a digital signal between its samples [1]. It maps each sample of the given signal to a translated and scaled copy of the Lanczos kernel, which is a sinc function windowed by the central hump of a dilated sinc function. The sum of these translated and scaled kernels is then evaluated at the desired points. Lanczos re sampling is typically used to increase the sampling rate of a digital signal, or to shift it by a fraction of the sampling interval. It is often used also for multivariate interpolation, for example to resize or rotate a digital image. Lanczos filter has been considered the "best compromise" among several simple filters. In image processing Lanczos window is applied to whole image first by rows and then by columns. There by sampling the required pixels and will reduce the size of the image.

The effect of each input sample on the interpolated values is defined by the filter's reconstruction kernel L(x), called the Lanczos kernel. It is the normalized sinc function sinc(x), windowed (multiplied) by the Lanczos window, or sinc window, which is the central lobe of a horizontally-stretched sinc function sinc(x/a) for $-a \le x \le a$.

$$L(x) = \begin{cases} sinc(x) sinc\left(\frac{x}{a}\right) & if -a < x < a \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Equivalently,

$$L(x) = \begin{cases} 1 & if\ x = 0 \\ \frac{a\sin(\pi x)\sin(\pi x/a)}{\pi^2 x^2} & if\ 0 < x < a \\ 0 & otherwise \end{cases} \quad (2)$$

The parameter a is a positive integer, typically 2 or 3, which determines the size of the kernel. The Lanczos kernel has $2a - 1$ lobes, a positive one at the center and $a - 1$ alternating negative and positive lobes on each side. Given a one-dimensional signal with samples S_i, for integer values of i the value S(x) interpolated at an arbitrary real argument x is obtained by the discrete convolution of those samples with the Lanczos kernel namely,

$$S(x) = \sum_{i=x-a+1}^{x+a} s_i\, L(x-i) \quad (3)$$

Where a is the filter size parameter and x is the floor function. The bounds of this sum are such that the kernel is zero outside of them. As long as the parameter a is a positive integer, the Lanczos kernel is continuous everywhere, and its derivative is defined and continuous everywhere (even at $x = \pm a$, where both sinc functions go to zero). Therefore, the reconstructed signal S(x) too will be continuous, with continuous derivative. The Lanczos kernel is zero at every integer argument x, except at $x = 0$, where it has value 1. Therefore, the reconstructed signal exactly interpolates the given samples: it has been deduced to $S(x) = S_i$ for every integer argument $x=i$.

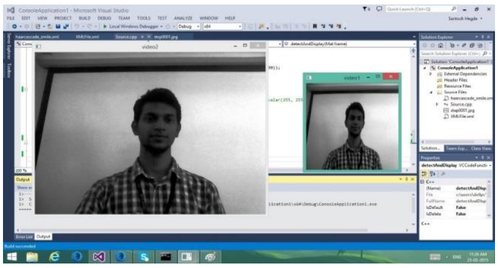

Fig.2. Image size reduced to 256x256 pixels.

The resizing of image is very important since, it helps to reduce the memory required to store the images. The Lanczos filter which is used to smoothly interpolate the value of a digital signal between its samples and it is much faster and less complex than others. Image size is reduced using Lanczos filter method as shown in figure 2.

B. Conversion to grey scale image

The input video taken from the camera is a sequence of colored frames. The intensity values of colored frames are divided into three colors namely red, green and blue. Grey scale image or frames contains the intensity value of each pixel. The spot at which intensity value is more that will be white spot and the spot at which the intensity value is minimum is dark spot. Calculation of co variance matrix needs a matrix of intensity value of pixels. Grey scale conversion has its own advantages; the most important is reduction in computations as greyscale images contains the single intensity values for each pixel it occupies less memory, and increase the speed. To convert a colored image to grey scale image many different algorithms such as lightness, average luminosity methods are available. The luminosity method is most sophisticated version of average method. It also averages the values, but it forms a weighted average to account for human perception. Human eyes are more sensitive to green than other colors, so green is weighted most heavily. The formula for intensity is

$$G_{luminous} = 0.21R+ + 0.72\ G + 0.07\ B \quad (3)$$

$G_{luminous}$ is the luminous of the grey scale. R, G and B are intensity values of each pixel respectively.

Fig.3. Conversion of color image to grey scale.

Grey scale images have single intensity value for each pixel and with this one can have single image matrix of intensity values. This matrix can be used for any further calculation. These matrix obtained have intensity values varying from 0 to 255 and one pixel is 8 bit. Grey scale conversion has many advantages. The most important is reduction in calculations, as greyscale images contains single intensity values for each pixel it occupies less memory, hence it increase in the processing speed. Figure 3 shows the conversion of the colored image to greyscale image, image in the left side is colored and the image shown in the right is the greyscale image.

C. Viola Jones method for face detection

Paul Viola and Michael Jones presented a quick and vigorous system for face identification which is 15times snappier than any method at the time of discharge with 95% exactness at around 17fps.

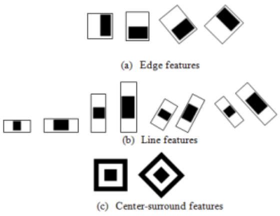

Fig.4. Different types of Haar features [2].

Figure 4 shows simple rectangular (Haar-like) features are reminiscent to Haar basis functions. These features are equivalent to intensity difference readings and are quite easy to compute. Described by Viola and Jones as a degenerative tree, and sometimes referred to as a decision stump, its use also speeds the detection process. A degenerative tree is the daisy chaining of general to specific classifiers, where by the first few classifiers are general enough to discount an image sub window and so on the time of further observations by the more specific classifiers down the chain, this can save a large degree of computation [4].

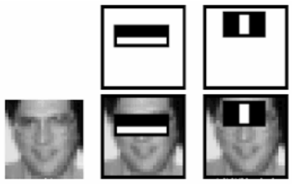

Fig.5. Different Haar like features applied on the human face [2].

Figure 5 shows the different Haar like features applied on the human faces. These features constitute a matrix which has values only 0 and 1. Block spot corresponds to 0 and white spot corresponds to 1. The speed with which features may be evaluated does not adequately compensate for their number.

For example, in a standard 24x24 pixel sub-window, there are a total of M=162,336 possible features when the Haar like features are applied in the human face as shown in the figure 5, and it would be prohibitively expensive to evaluate them all when testing an image. Thus, the object detection framework employs a variant of the learning algorithm AdaBoost to both select the best features and to train classifiers that use them. AdaBoost is used to remove all the irrelevant features and employs only relevant features. Therefore by selecting only relevant features the time involved in applying all features to every input image can be reduced considerably. This algorithm constructs a "strong" classifier as a linear combination of weighted simple "weak" classifiers.

$$h(x) = sign\left(\sum_{j=1}^{M} \alpha j h j(x)\right) \quad (4)$$

Each weak classifier is a threshold function based on the feature f_j.

$$hj(x) = \begin{cases} -sj & if\ fj < \theta j \\ sj & otherwise \end{cases} \quad (5)$$

The threshold value $\theta_j$and the polarity $s_j \in \pm 1$ are determined in the training, as well as the coefficients α_j.

Input: Set of N positive and negative training images with their labels(x^i, y^i). If image i is a face,$y^i = 1$ if not, $y^i = -1$.

1. Initialization: assign a weight $w_1^i = \frac{1}{N}$to each image i.
2. For each feature $f_j$with j=1,…,M.
   a. Renormalize the weights such that they sum to one.
   b. Apply the feature to each image in the training set, then find the optimal threshold and polarity $\theta_j s_j$that minimizes the weighted classification error.
   c. Assign a weight $\alpha_j$to h_j that is inversely proportional to the error rate. In this way best classifiers are considered more.
   d. The weights for the next iteration, i.e. w_{j+1}^i, are reduced for the images i that were correctly classified.
3. Set the final classifier to

$$h(x) = sign\left(\sum_{j=1}^{M} \alpha_j h_j(x)\right) \quad (6)$$

The key advantage of a Haar-like feature over most other features is its calculation speed. Due to the use of integral images, a Haar-like feature of any size can be calculated in constant time. A simple rectangular Haar-like feature can be defined as the difference of the sum of pixels of areas inside the rectangle, which can be at any position and scale within the original image. This modified feature set is called 2-rectangle feature. Viola and Jones also defined 3-rectangle features and 4-rectangle features. The values indicate certain characteristics of a particular area of the image. Each feature type can indicate the existence (or absence) of certain characteristics in the image, such as edges or changes in texture. For example, a 2-rectangle feature can indicate where the border lies between a dark region and a light region.

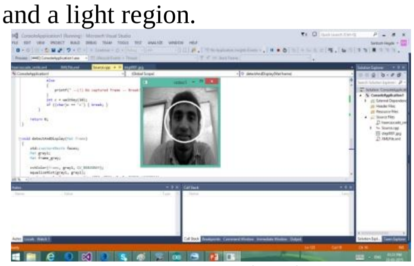

Fig.6. Detection of the face

Figure 6 shows the face detection using OpenCV. The technique used is Viola Jones algorithm which consists of steps like integral image whose main purpose is to reduce the complexity in calculating the sum of pixels of a particular region and Haar cascade classifiers which consists of many weak classifiers and each of these weak classifiers has a threshold value and based on this the presence or absence of face in any given image is detected.

III. FEATURE EXTRACTION AND EMOTION CLASSIFICATION

This section describes the features extracted from the face i.e. calculation of the variance co variance matrix from which Eigen value are calculated in order to classify the expressions. It also describes the database used for training the images i.e.

Japanese female facial expression (JAFFE). The images of these databases are trained and classification algorithms are applied to classify the expressions.

A. *Covariance matrix*

Variance and Covariance are a measure of the "spread" of a set of points around their center of mass (mean). Variance is a measure of the deviation from the mean for points in one dimension e.g. heights. Covariance as a measure of how much each of the dimensions varies from the mean with respect to each other. Covariance is measured between 2 dimensions to see if there is a relationship between the 2 dimensions. The covariance between one dimension and itself is the variance.

B. *Variance*

Variance is a measure of the variability or spread in a set of data. Mathematically, it is the average squared deviation from the mean score.

$$Var(X) = \Sigma (Xi - X)^2 / N = \Sigma x_i^2 / N \quad (7)$$

N is the number of scores in a set of scores.
X is the mean of the N scores.
X_i is the i^{th} raw score in the set of scores.
x_i is the i^{th} deviation score in the set of scores.
Var(X) is the variance of all the scores in the set

C. *Covariance*

Covariance is a measure of the extent to which corresponding elements from two sets of ordered data move in the same direction.

$$Cov(X, Y) = \Sigma (Xi - X) (Yi - Y) / N = \Sigma x_i y_i / N \quad (8)$$

N is the number of scores in each set of data
X is the mean of the N scores in the first data set
X_i is the i^{th}raw score in the first set of scores
x_i is the i^{th} deviation score in the first set of scores
Y is the mean of the N scores in the second data set
Y_i is the i^{th} raw score in the second set of scores
y_i is the i^{th} deviation score in the second set of scores
Cov(X, Y) is the covariance of corresponding scores in the two sets of data

D. *Variance-covariance matrix*

Variance and covariance are often displayed together in a variance-covariance matrix. The variances appear along the diagonal and covariance appears in the off-diagonal elements.

$$V = \begin{bmatrix} \Sigma x_1^2 / N & \Sigma x_1 x_2 / N & \dots & \Sigma x_1 x_c / N \\ \Sigma x_2 x_1 / N & \Sigma x_2^2 / N & \dots & \Sigma x_2 x_c / N \\ \dots & \dots & \dots & \dots \\ \Sigma x_c x_1 / N & \Sigma x_c x_2 / N & \dots & \Sigma x_c^2 / N \end{bmatrix}$$

V is a *c* x *c* variance-covariance matrix.
N is the number of scores in each of the *c* data sets.
x_i is a deviation score from the *ith* data set.
$\Sigma x_i^2 / N$ is the variance of elements from the *i*th data set.
$\Sigma x_i x_j / N$ is the covariance for elements from the *i*th and *j*th data sets.
Suppose **X** is an *n* x *k* matrix holding ordered sets of raw data.

E. *Calculation of Eigen values*

Emotions are classified based on the dominant Eigen values. For each emotion, a range of Eigen values are calculated and stored for training images. For an input image, the dominant Eigen value is calculated and depending on the range of Eigen values of training image it falls into, the emotions are detected. Eigen values for the training images are calculated in the beginning and with that dominant Eigen value for each image is determined. For each expression a range of Eigen value is defined. The Eigen values for each video frame are calculated from the covariance matrix obtained from the input video frame.

The Eigen values of a matrix I can be determined by finding the roots of the characteristic polynomial. Explicit algebraic formulas for the roots of a polynomial exist only if the degree n is 4 or less. According to the Abel Ruffini theorem there is no general, explicit and exact algebraic formula for the roots of a polynomial with degree 5 or more.

Let I be an covariance matrix of order n, λ is an Eigen value if and only if there exists a non-zero vector *C* such that

$$A\,C = \lambda\,C$$

$$(A-\lambda I_n)C=0 \quad (9)$$

Where I_n is the identity matrix of order n.

This is a linear system for which the matrix coefficient is $A-\lambda I_n$. This system has one solution if and only if the matrix coefficient is invertible, i.e. $\det(A-\lambda I_n)\neq 0$. Since, the zero-vector is a solution and *C* is not the zero vector, then it must be

$$\det(A-\lambda I_n)=0 \quad (10)$$

This equation is evaluated to find the value of λ, which is the desired Eigen value.

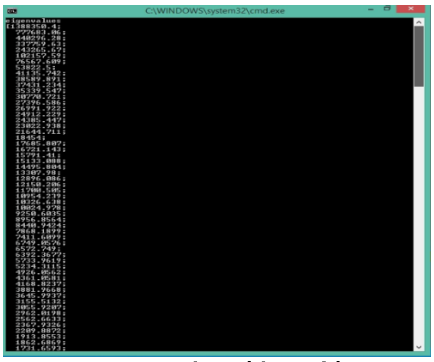

Fig.7. Eigen values of detected face

Fig 7 shows the Eigen values obtained for the detected face. The Eigen values of matrix obtained directly from the image can be determined by finding roots of characteristic polynomial. Explicit algebraic formulas for roots of a polynomial exist only if degree n is 4 or less. The Eigen values are stored for each image in memory. For any given image its Eigen values are compared with pre stored values and one with closest approximation is chosen and its corresponding expression is accessed from memory.

F. *JAFFE database*

JAFFE database contains 213 images of 7 facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Each image has been rated on 6 emotion adjectives by 60 Japanese subjects. The database was planned and assembled by Michael Lyons, Miyuki Kamachi, and Jiro Gyoba. The JAFFE has been chosen since, for each emotion the database consists of many images there by the efficiency for classification can be improved.

G. Training images

A database of images of people for different expressions is developed and is used as reference for classification of emotions. A training process for facial expression recognition is usually performed sequentially in three stages: feature learning, feature selection, and classifier construction. First, features that capture expression related facial appearance/geometry changes are extracted from images or video sequences. These features can be either hand-designed or learned from training images. Then, a subset of the extracted features is the most effective to distinguish.

Fig.8. Samples of image database

Figure 8 shows the samples of image database used. The database used here is the Japanese database in which for each and every expressions such as angry, happy, and sad, many images of different people are considered since each person will have a different expression for a given emotion. The given image is compared with these images in the database and the corresponding expression which would already be stored in the memory would be read out.

H. Emotion classification

Based on the range of Eigen values calculated between each training image and image captured from the camera, emotion are determined.

Table .1. Emotion classification range

	EIGEN VALUE RANGE	
EXPESSION	**FROM**	**TO**
NEUTRAL	2.512	2.835
SAD	3.02	3.599
SMILE	3.701	3.951
ANGER	4.021	4.989
SURPRISE	5.112	5.686

Five expressions have been detected and classified such as neutral, sad, smile, anger and surprise. For classification of these images, 30 training images were evaluated and the range of Eigen values was determined in table 1.

I. Emotion classification using clustering

Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters). It is a task of exploratory data mining, and a common technique for statistical data analysis used in many fields including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics.

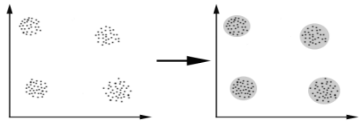

Fig.9. Group of 4 clusters [3]

Figure 9 shows group of clusters, cluster is a collection of objects which are *similar* between them and are *dissimilar* to the objects belonging to other clusters.

J. K-Means clustering

K-means algorithm is most popular partitioning based clustering technique [6]. It is an unsupervised algorithm which is used in clustering. It chooses centroid smartly and it compares centroid with the data points based on their intensity and characteristics and finds the distance, data points which are similar to centroid are assigned to the cluster having centroid. New k centroids are calculated and thus k-clusters are formed by finding data points nearest to the clusters. K-means algorithm is a well-known clustering algorithm popularly known as Hard C Means algorithm. This algorithm splits the given image into different clusters of pixels in the feature space, each of them defined by its centre. Initially each pixel in the image is allocated to the nearest cluster. Then the new centers are computed with the new clusters. These steps are repeated until convergence. It is required to determine the number of clusters K first. Then the centroid will be assumed for these clusters. It could be assumed that random objects as the initial centroids or the first K objects in sequence could also serve as the initial centroids. Efficiency is determined by the number of centers assigned by the user. From the centers assigned, the range of the centre is calculated from distance formula. Sum of all the objects in a particular cluster is added and mean is calculated which becomes a new centroid of the cluster. This is continued until the centroids remain unchanged.

IV. RESULTS

The experiment was carried out on 25 different people and it was tested on different background for 5 different expressions such as Neutral, Sad, Smile, Angry and Surprise. The program was executed and obtained results are tabulated in each case. Camera takes 30 frames per second which is then processed at a rate of 17 frames per second. For classification both K-means algorithm and Eigen value based method also called thresholding methods are used.

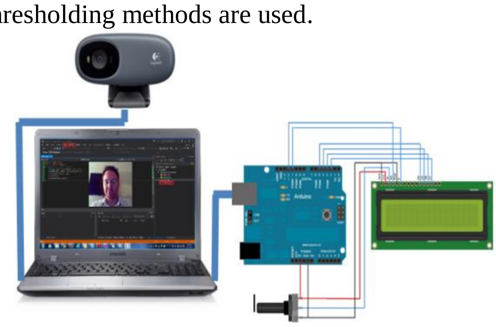

Fig.10. Experimental setup of emotion detection and classification.

Figure 10 shows the experimental setup of emotion detection and classification. Input is taken from the Logitech web camera, input video is processed using OpenCV. It consists of ATmega328 microcontroller and 16X2 Liquid crystal display.

A. *Smile output*

Fig.11. Detected emotion SMILE

A frame of the video is captured and given as an input to the face detection algorithm. Face is detected and positions of the face are returned and an ellipse is inserted to the positions of a face indicating that face is detected. If the range of Eigen values tabulated is between 3.701 and 3.951 then, the expression is detected as smile. Figure 11 shows the expression detection (SMILE) of three persons. The window on the right side in the figure displays the detected expression. If no expression is shown the "not detected" is displayed. The expression for the current iteration is highlighted in red line.

B. *Anger output*

Fig.12. Detected emotion ANGER

C. *Surprise output*

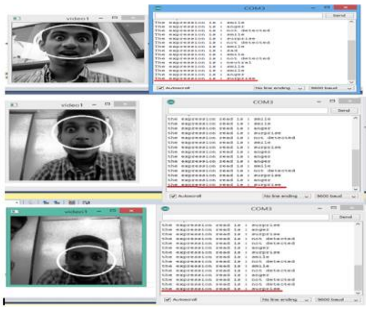

Fig.13. *Detected emotion SURPRISE*

D. *Neutral output*

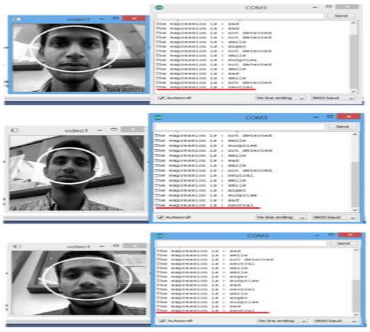

Fig.14. *Detected emotion NEUTRAL*

E. *Sad output*

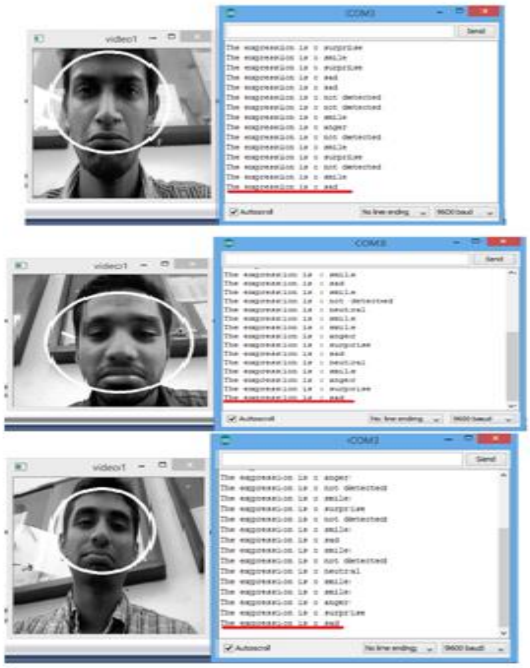

Fig.15. *Detected emotion SAD*

Figure 12-15 shows the expression detection of three persons for anger, surprise; neutral, sad and range of Eigen values tabulated are 4.021 and 4.989, 5.112 and 5.686, 2.512 and 3.825, .02 and 3.889 respectively.

OUTPUT EFFICIENCY

A. Output efficiency of Eigen value (or) thresholding classification

Table 2. Output efficiency of five expressions using thresholding classification.

EXPRESSI ON	NO. IMAGES CONSIDERED	DETECTED CORRECT IMAGES	DETECTED WONG IMAGES	EFFECIENCY
HAPPY	20	18	2	90%
ANGRY	20	17	3	83%
SURPRISE	20	16	4	80%
SAD	20	16	4	80%
NEUTRAL	20	18	2	90%

Each expression was evaluated by taking expressions of 20 persons and the output efficiency of each expression is tabulated as shown in table 2.

B. *Output efficiency of K-means classification*

Table. 3. Output efficiency of five expressions using K-means classification.

EXPRESSION	NO. IMAGES CONSIDERED	DETECTED CORRECT IMAGES	DETECTED WONG IMAGES	EFFECIENCY
HAPPY	20	16	4	75%
ANGRY	20	10	10	50%
SURPRISE	20	15	5	75%
SAD	20	14	6	66%
NEUTRAL	20	15	5	75%

Each expression was evaluated by taking expressions of 20 persons and the output efficiency of each expression is tabulated as shown in table 3.

COMPARISON OF CLASSIFICATION METHODS

Table. 4. Comparison of K-means classification and thresholding classification

EXPRESSION	K-MEANS ALGORITHM EFFICIENCY	THRESOLDING ALGORITHM EFFICIENCY
HAPPY	75%	90%
ANGRY	50%	83%
SURPRISE	75%	75%
SAD	66%	80%
NEUTRAL	75%	90%

Table 4 shows the comparison of K-means classification and thresholding classification. From the comparison table it can be inferred that classification through Thresholding algorithm yields better efficiency.

V. CONCLUSION

The emotion detector system for detecting expressions had been developed. For face detection viola and jones algorithm has been used. For classification of emotions, thresholding classification has been proposed and this algorithm proves to be much more efficient than the K-means clustering algorithm. Module is developed and tested for different emotions. Classification algorithm called Thresholding which uses Eigen value range. For a particular Eigen value range, the expression is obtained satisfactorily and the efficiency of facial expression detection for smile, anger, surprise, sad and neutral are 90%, 83%, 80%, 80% and 90% respectively. Thresholding algorithm achieves higher efficiency than the K-means clustering. Hence it is highly accurate and reliable detection of emotions which can aid autism patients.

Currently we are building a decision system to notify the expression and enhancing the proposed thresholding algorithm for lower contrast i.e. with lesser lighting, increasing the precision of the Eigen values to increase the efficiency; further thresholding is applying for other expression such as disgust, hatred, fearful.

References

[1] Yisu Zhao “Part-based PCA for facial feature extraction and classification”, *Haptic Audio visual Environments and Games IEEE International Workshop*, 2009, pp. 99 - 104.

[2] Zhengmao Ye, Mohamadian, H. and Yongmao Ye, “Comparative study of linear and nonlinear color model identification based optimal feature extraction”, *IEEE 12th International Symposium on Computational Intelligence and Informatics*, Budapest, 2011, pp.97 – 102.

[3] Zahir, N.B., Samad, R. and Mustafa, M. “Initial experimental results of real-time variant pose face detection and tracking system”, *IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Melaka, 2013, pp.264 – 268

[4] Agrawal, Samiksha Khatri and Pallavi, “Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis”,*5th International Conference on Advanced Computing & Communication Technologies (ACCT)*, Haryana India, 2015, pp. 108 – 112.

[5] Yisu Zhao “Part-based PCA for facial feature extraction and classification”,*Haptic Audio visual Environments and Games IEEE International Workshop*, 2009, pp. 99 - 104.

[6] Selim, Shokri Z “K-Means-Type Algorithms: A Generalized Convergence Theorem and Characterization of Local Optimality”, *IEEE Transactions on Pattern Analysis and Machine Intelligence* ,pp. 81 – 87.

[7] F. Happe, J. Briskman, “Exploring the cognitive phenotype of Autism: weak central coherence in parents and siblings of children with autistic, Experimental tests”, *Journal of Child Physiology and Psychiatry and Allied Disciplines*, vol. 42, no. 3, 2011, pp: 299-307.

[8] H. V. Ravish Aradhya, Mohana, “*Real Time Objects Detection and Positioning in Multiple regions Using Single Fixed Camera View for Video Surveillance Applications”* Proceedings of the 2015 IEEE International Conference on ‘Electrical, Electronics, Signals, Communication & Optimization(EESCO-2015) ,vizac, pp.1-6, 24-25 January 2015

[9] H. V. Ravish Aradhya, Mohana, *“Real-Time Implementation of object detection and Tracking on DSP for video survillance applications”* Proceedings of the 2016 IEEE International Conference on ‘Recent Trends in Electronics, Information & Communication Technolog’ pp.1-5, may 18-20 2016, Bangalore, IEEE 2016.

[10] H. V. Ravish Aradhya, Mohana, *“Real-Time Implementation of object detection and Tracking on FPGA for video survillance applications”* 2016 IEEE International Conference on ‘Distributed Computing,VLSI, Electrical Circuits and Robotics(DISCOVER)’ National Institute of Technology surathkal, Karnataka, NITK Surathkal 13-14 August 2016

[11] H. V. Ravish Aradhya, Mohana, *“Elegent and Efficient Algorithms for real time object detection counting and classification for video survillance applications”* Proceedings of 2016 IEEE 2nd International Conference on ‘Circuits, control, Communication and computing (I^4C)’ Bangalore, Karnataka, MSRIT, 5-6 October 2016.