iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** III **Month of publication:** March 2018

DOI: http://doi.org/10.22214/ijraset.2018.3551

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

Recognition of an Emotion using Principle Component Analysis of Hand Gesture of a Human

Mukesh T. Patel[1]
[1]*Department of Mathematics, Uka Tarsadia University-Bardoli*

Abstract: This paper explores recognition of human emotion through the hand gestures generated for particular emotion. In non-verbal reflection of an emotion hand generates various gestures using its most flexible parts viz. Fingers, Palm and Arm. In recognition process, an image or sequence of images of a hand gesture is first normalized before extracting the features from them. Thereafter, as a feature of these normalized gesture images, the Principle Components (PCs) of images are analysed. Furthermore, Neural Networks (NN), using Back Propagation Algorithm classifies the obtained PCs and finally the input gesture of an emotion is matched into the classes of pre stored trained hand gesture image sets of seven basic emotions Neutral, Happy, Sad, Fear, Anger, Disgust and Surprise.
***Keywords*: Emotion, Hand Gesture, PCA, Neural Network, Back Propagation.**

I. INTRODUCTION

During Communication a human expresses his thoughts and views by both verbally and non-verbally. In which, non-verbal communication performs its major role as compare to verbal to convey message more efficiently [3], [4]. Although, a whole human body is always remained an interactive part during communication but in the creation of gesture or posture related to an emotion, the maximum contribution is of its most flexible parts Head, Hands and Legs [1], [2]. In this paper, only the hand gestures is being considered that are further classified in the gestures generated by Fingers, Palm and Arm either individually or in-group with respect to an emotion [3], [4], [5]. However, this non-verbal hand gestures reflect emotion physically so that it can be captured in the form of images and can be analysed with the knowledge of image processing [6] as a computational transformation of mathematics to reveal the felt emotion of a human being. The proposed, Principle Components Analysis of Hand Gesture based Mathematical System works in three phases on captured images of hand gestures (i) Image Normalization (ii) Feature Extraction- Principle Component Analysis (PCA) and (iii) Classification by Neural Network (NN) using Back Propagation Algorithm [9]. Which classifies the input hand gestures of an emotion into the pre-stored classes of seven basic emotions mentioned earlier.

II. PRINCIPLE COMPONENT ANALYSIS OF HAND GESTURE BASED MATHEMATICAL SYSTEM

The MATLAB architecture to recognize an emotion of a human being from hand gestures is shown in Fig. 1, where the image or sequence of image (video) of a hand gesture is an input and the related emotion will be the output of the system. The different phases of the system is described as under.

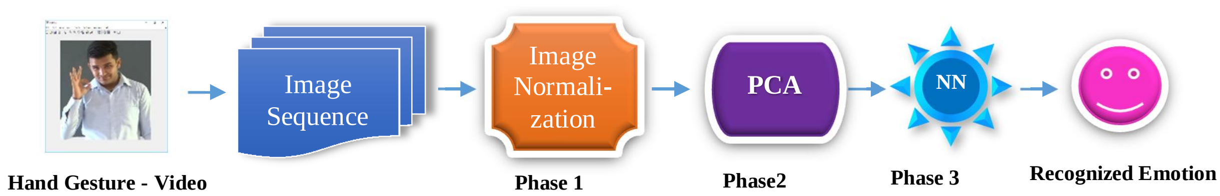


Fig. 1 Principle Component Analysis of Hand Gesture based Mathematical System

Initially the input video of hand gesture is converted in to the image sequence (frames) and consider only those frames, which are significantly variant in the form of appearance of gestures related to an emotion.

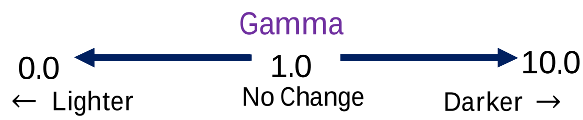
A. Phase-1: Image Normalization

To enhance the quality of features, the Images are normalized by the following ways.

1) *Gamma Correction (on RGB Image) [6]:* Gamma correction improve the visibility of an image by balancing darkness and lightness of an image. Gamma function: For the original image X the Gamma corrected new image is $X_{NEW} = X^{GAMMA}$

$$\text{Where Gamma} = (I_{out})/(I_{in}),\ I_{in} = \text{Image input and } I_{out} = \text{Image output.}$$

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com



Gamma ∊ [0.0, 10.0], Where, 0.0 indicates lightness and 10.0 indicates darkness while 1.0 for no change of image in MATLAB [10].

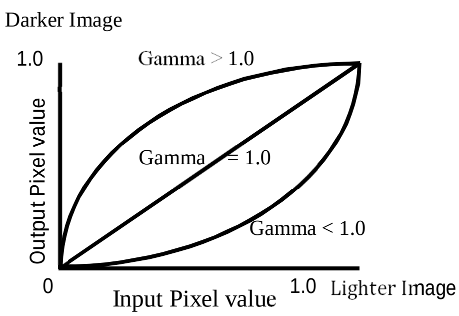


Fig. 2 Gamma Curve

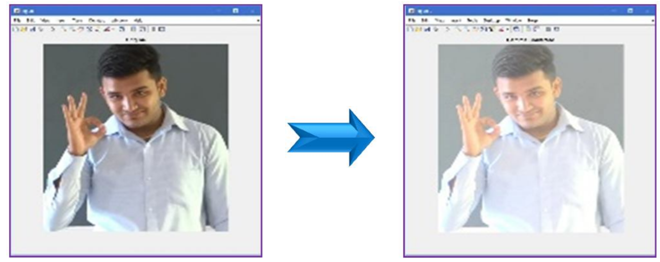


Fig. 3 Gamma Corrective Image

Fig. 2 shows several gamma curves demonstrating the effect of gamma value has on the shape of the gamma curve.

2) *RGB to Gray conversion [6], [10]:* The coloured image has three dimensional pixel values (Matrix) Red, Green and Blue (RGB).

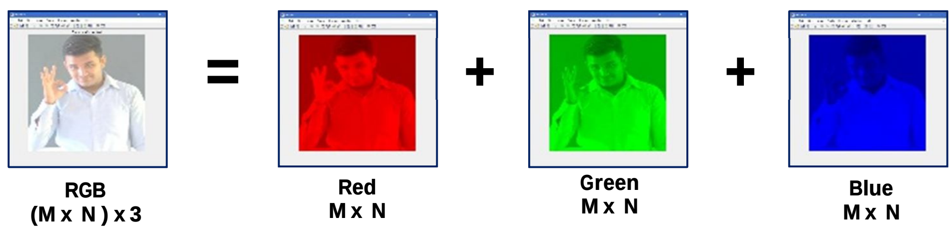


Fig. 4 RGB Components of an Image

Moreover, PCA is irrelevant to colour such that it is not necessary to process all three matrices RGB, in place of it is better to convert these three matrices in a single Gray matrix by the standard relation [6];

$$I_{(Gray)} = \mathbf{0.2989 * R + 0.5870 * G + 0.01140 * B}$$

A gray image has a single pixel value which lies between [0, 255] as per the intensity of gray value of a pixel [10].

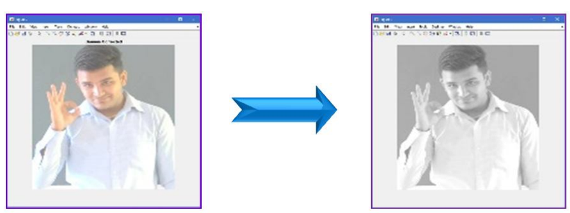


Fig. 5 RGB→ Gray Image Conversion

3) *Histogram Equalization [6], [10]:* In a feature extraction, high contrast gray image makes each pixel's gray value significantly differ from each other that increase the quality of features. Intensity levels in the range [0, *G*] is defined as the discrete function; $h(r_k) = n_k$, Where r_k is the k^{th} intensity level in the interval [0, G] and n_k is the number of pixels in the image whose intensity level is r_k and G = 255. Therefore, the normalized histograms function is;

$$p(r_k) = \frac{h(r_k)}{n} = \frac{n_k}{n}$$

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

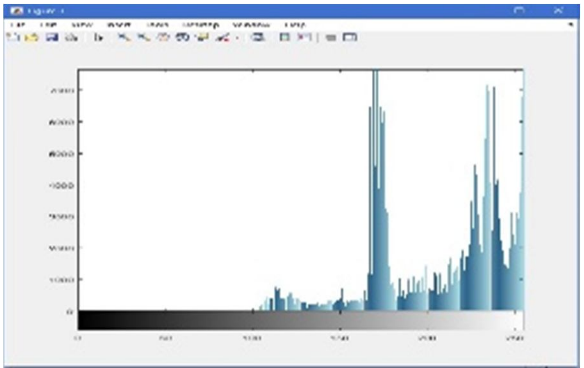


Fig. 6 Histogram of Original Gray Image Diagram

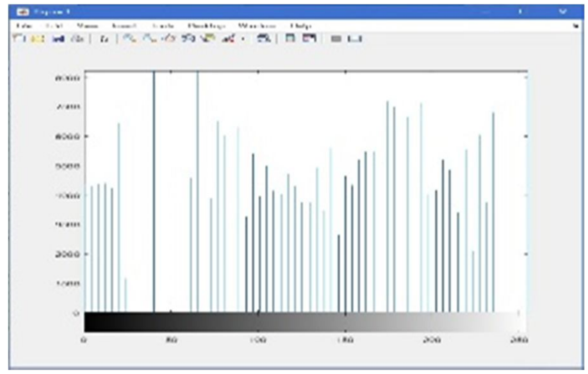
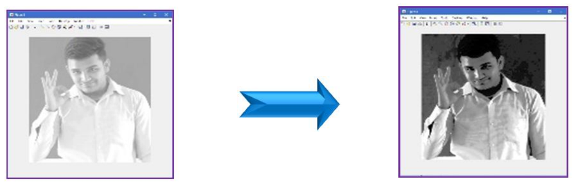


Fig. 7 Histogram Equalized Gray Image Diagram



Original Gray Image

Histogram Equalized Gray Image

Fig. 8 Histogram Equalization

4) Image resizing [10]

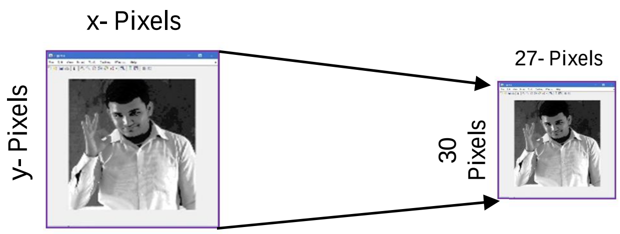


Fig. 9 Resizing of Histogram Equalized Gray Image

Resizing image is a customized part of normalization to make the recognition process smoother and faster. The size of image should not be so big, which increases the process time and not so small, which is not sufficient for feature extraction such that, the moderate size should be chosen that does not affect the recognition process [8]. Fig. 9 shows the resized normalized image is now ready for Phase – 2, feature extraction.

B. Phase - 2 : Feature Extraction - PCA [7]:

Consider the size of input image of the dimension p × q pixels, by transforming it in a column vector prepare a matrix X of all N hand gesture images of an emotion in order (p × q, 1). Let each column vector of the matrix X is denoted by x_j, where $j = 1, 2, 3 \ldots n$. Find the mean of corresponding row vector, which gives an average pixel of similar gesture.

$$\text{Mean } \bar{X} = \frac{1}{n}\left(\textstyle\sum_{j=1}^{n} x_{ij}\right), \text{ Where i = 1, 2, 3} \ldots (p \times q).$$

Thus, the normalized column vector of a given matrix X is obtained by $\sigma_j = x_j - \bar{X}$. Find n orthonormal vectors U_n, represents the best distribution of the data. The kth vector U_k, is chosen such that

$$\lambda_k = \frac{1}{n}\textstyle\sum_{j=1}^{n}\left(U_k{}^T \sigma_j\right)^2 \quad (1)$$

Which is maximum, where $U_j{}^T U_k = \begin{cases} 1 \; if \; j=k \\ 0 \; other\,wise \end{cases}$. Extend (1) by multiply $U_k{}^T$ to both sides of the equation,

$$\lambda_k U_k{}^T = U_k{}^T \frac{1}{n}\sum_{j=1}^{n}\left(\sigma_j \, \sigma_j{}^T\right) U_k U_k{}^T$$

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

Assume the covariance matrix, $C = \phi\, \phi^T$, Where $\phi = [\sigma_1, \sigma_2, \sigma_3, \dots . \sigma_n]$, We obtain

$$\lambda_k \mathrm{U}_k^{\ T} = \mathrm{U}_k^{\ T} C \qquad (2)$$

Transpose both the side of (2), we have $\lambda_k (\mathrm{U}_k^{\ T})^T = C^T (\mathrm{U}_k^{\ T})^T = C^T \mathrm{U}_k$ and $C^T = \phi\, \phi^T = C$, Then we can conclude that $\lambda_k \mathrm{U}_k$ = C U_k. So the vectors U_k and scalars λ_k are the eigenvectors and eigenvalues respectively, of the covariance matrix $C = \phi\, \phi^T$.
Because the size of covariance matrix C is $pq \times pq$, it is time-consuming to determine pq eigenvalues and eigenvectors. It is necessary to reduce the complexity of the computation. Define a matrix C' = A^T A , such that size of C′ is $N \times N$.
Let, v_i is the eigenvectors of the matrix C′, we have $A^T A v_i = \eta_i v_i$. Multiply A to both sides, we get

$$A A^T A v_i = \eta_i A v_i \qquad (3)$$

From (3) that Av_i is the eigenvector of $C = AA^T$. So we can now reduce the method by derive r $(r < pq)$ eigenvectors (v_i) of the matrix C′ and derive the r eigenvectors of covariance matrix C by multiplying A to v_i.
The hand gesture images are represented by projecting the data in the image space onto the hand gesture space which reduces the dimensions as well. By using eigen images, the hand gesture images can be transformed by simple projection operation as,

$$y_i = U_i^{\ T} (\mathrm{x} - \bar{X}), \quad i = 1, 2, 3, \dots\ r. \qquad (4)$$

Where, y_i is projection of hand gesture image. Thus the Projection vector $Y^T = [y_1, y_2, \dots, y_p]$. The projection vector Y will be treated as the features for further training process.
There may be similarity in the obtained eigenvectors of input images which is unnecessary increase the time period of classification even consumes much memory of the system. Thus, we will consider only those eigenvectors which is significantly differ from each other, consist all the characteristics of input images. To choose those significantly variant eigenvectors, all vectors are plotted in the decreasing order of their respective eigenvalues can be seen in Fig. 10.

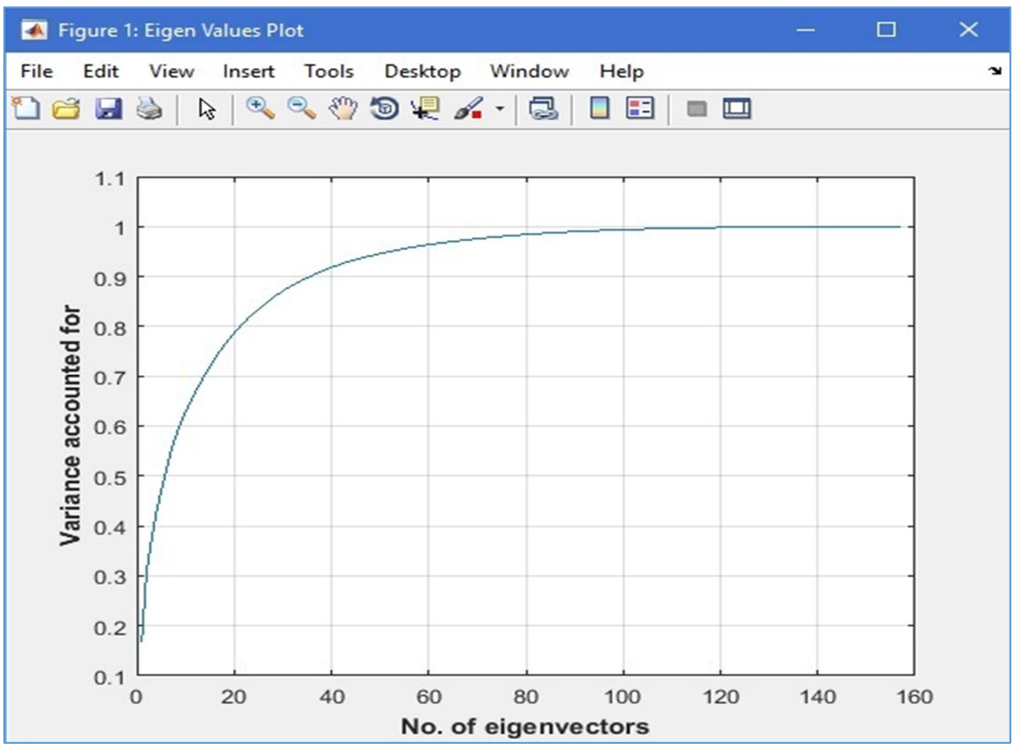


Fig. 10 Graph of Eigen Vectors (Images) Based on Respective Decrease Ordered Eigen Values

It is clearly seen in Fig. 10, total 160 eigenvectors of input images are plotted in the decreasing order of their eigenvalues such that nearby 40 to 50 eigenvectors are only significantly vary from each other and thereafter remaining are constants means, remaining eigenvectors have same characteristics of input images which are irrelevant in the classification, that are not taken into the consideration.
Using this eigenvectors the feature vectors of original hand gesture images referred as eigen images obtained by PCA is projected in Fig. 11.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

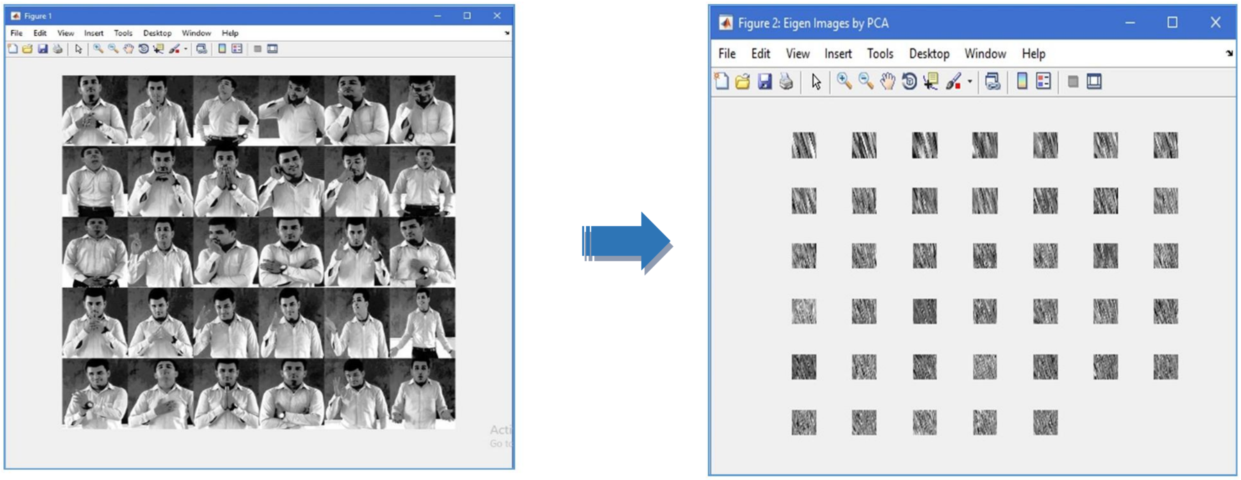


Fig. 11 Eigen Images by Principle Component Analysis

C. Classification

At phase – 2 obtained PCA based eigen images of hand gesture with respect to emotion is classified at this level using neural network (NN) [8], [9]. In phase – 3 NN is trained to classify PCA eigen images by, Multilayer Perceptron (MLP) with Back Propagation classification algorithm [9]. Also, the pre required parameter to train NN successfully are obtained before training viz. Error Surface, Number of Hidden Layers, Learning Rate, Momentum, Input Standardization & Weight Initialization, Training Stopping Criteria etc.

Multilayer Perceptron (MLP) is widely used neural network having many application especially for the classification pattern. MLPs imply feed-forward networks and Back- Propagation Algorithm [9].

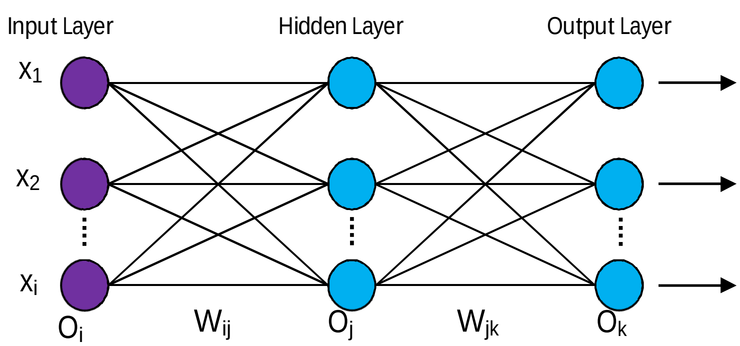


Fig. 12 Fully Connected, Feed-Forward MLP Network

Multilayer Perceptron Neural Network includes three different layers Input layer, Hidden Layer and Output layer. The whole training sample $X = x_1, x_2, \ldots, x_i$, is considered at input layer with i input nodes and outputs are set to output layer having k output nodes. Input nodes are connected to output nodes via j hidden layers. Also, Weight W_{ij} indicates the weight form a unit j in one layer to a unit i in the previous layer. The sum of all weights are ultimate output of input nodes. Input nodes is correspond to pixel values of PCA based eigen images and output layer corresponds to classes of emotions into the database. Here, output layer is trained to respond +1 for matching and 0 for others. But, in real practice outputs are vary between 0 and +1.

The NN algorithm can be converted to the standard gradient descent version of BACKPROPAGATION if the gradient becomes:

$$\delta_k = \sum_{pattern}^{n} O_{n.k}\,(1 - O_{n.k})(t_{n.k} - O_{n.k}) \tag{5}$$

Where, n is the number of the training patterns and k is the number of output units. Usually δ_k is divided by the total number of training patterns in order to constrain the weight update to the mean of the updates caused by each training pattern.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

TABLE I
BACK PROPAGATION ALGORITHM

✓ Initialization of network weights
✓ Until the termination**:**
{For each training example:
{Propagate input is forwarded to the network and compute the observed output.
Propagate the errors backward as follows:
For each network output unit k, calculate its error term $\delta_k = O_k(1-O_k)(t_k-O_k)$
For each hidden unit calculate its error term $\delta_h = o_h(1-o_h)\sum_{k\ outputs} W_{kh}\delta_k$
Finally, update each weight $w_{ji} = w_{ji} + \Delta w_{ji}$,
Where $\Delta w_{ji} = -\eta.\delta_j x_{ji}$ }
}
(Here, 'ji' indicates from unit 'i' to 'j')

1) *Error Surface* [8], [9]: The gradient of the error function for the output 'k'. The total error, in standard gradient descent version of BACKPROPAGATION is the SSE (Sum of Squared Errors):

$$E(\overline{w}) = \sum_{n\ pattern}\sum_{k\ outputs}(t_{n,k} - o_{n,k})^2 \tag{6}$$

Where E is a function of the network's weight vector. Gradient descent starts with an arbitrary weight vector and tries to minimize E at each step. In order to go deeper in these multidimensional surface weights must be updated in the direction of the negative of the gradient $\frac{dE(\overline{w})}{d(\overline{w})}$.

2) *Size of Hidden Layer* [8], [9]: For Pattern recognition one hidden layer is standard and sufficient so that we have used one-hidden layer.
3) *Number of neurons at hidden layer* [8], [9]*:* To choose optimum number of neurons in hidden layer, prior training or testing is still an issue of research, it is totally based on experiment and on the choice of other parameters of NN.
4) *Learning Rate* [8], [9]: In Table I weights are updated by $\Delta w_{ji} = -\eta.\delta_j x_{ji}$. This is called learning rate of the Back propagation algorithm. The performance of the algorithm is very sensitive to the proper setting of the learning rate. If the learning rate is set too high then, the algorithm may oscillate and become unstable. If the learning rate is too small then, the algorithm will take too long to converge. It is not practical to determine the optimal setting for the learning rate before training. This is most likely obtained by trial and error.
5) *Momentum Term* [8], [9]: A common modification of the basic weight update rule is the addition of a momentum term. By adding this term to the formula of the final step in Table I, the updated rule will be: $\Delta W_{ji} = \eta\delta_j x_{ji} + \alpha\Delta W_{ji}(n-1)$. Therefore, the update in iteration is affected by the update in n^{th} iteration multiplied by a factor 'α', called momentum. Momentum takes values in the range of $0 \le \alpha < 1$. Empirical evidence shows that the use of a momentum in the BACKPROPAGATION algorithm can be helpful in speeding the convergence and avoiding local minima in the error surface.
6) *Input Standardization and Weights Initialization* [8], [9]*:* In the classification process of NN, at hidden nodes every weights are set to zero and at output nodes it is set to random which gives better appearance of hand gesture than the other settings.
7) *Training Stopping Criteria* [8], [9]*:* There are many options to set a criteria of stopping training process viz. Fixed number of iterations (Epochs) to be performed, Use particular threshold for error, Use threshold for error gradient, Early stopping with validation error.

III.EXPERIMENTAL RESULTS

In this research work, we are identifying a single emotion viz. Neutral, Happy, Anger, Sad, Disgust, Fear and Surprise using respective hand gestures of 10 individuals. As it is discussed earlier, a hand gesture is divided in three gesture categories, gestures created by fingers, palm and arm such that each and every gestures related to an emotion is classified accordingly.

The variety of gestures of fingers, palm and arm related to individual single emotion is stored in an image database. The database further divided into two different sets Training and Testing sets. Here, the size of training data set is taken as 90% of all images and

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

10% size of testing set. Before, applying rather simulating proposed “Principle Component Analysis of Hand Gesture based Mathematical System” on our database some basic parameters of NN discussed in previous section have to set in prior.
The optimum topology for performing Neural Network based classification of PCA’s of hand gestures to recognize a single emotion for an individual human being is as follows.

TABLE II
OPTIMUM TOPOLOGY OF NEURAL NETWORK FOR HAND GESTURE – SINGLE EMOTION

Parameter	Expression
Data set	Hand Gesture blocks 27 x 30 pixels
Training set	90 %
Test set	10 %
Validation set	0%
Input Standardization	PCA40
Weight Initialization	Hidden Layer = 0 , Output Layer = random
Training Algorithm	Gradient Descent with Momentum
Transfer Function	Both layer use log-sigmoid
Neurons in Hidden Layer	26 neurons
Learning rate	1.4
Momentum term	0.6
Minimum training MSE	0.014288
Maximum Epochs	1600

With the consideration of above topology the performance output of the proposed system in MATLAB is seen in Fig. 13.

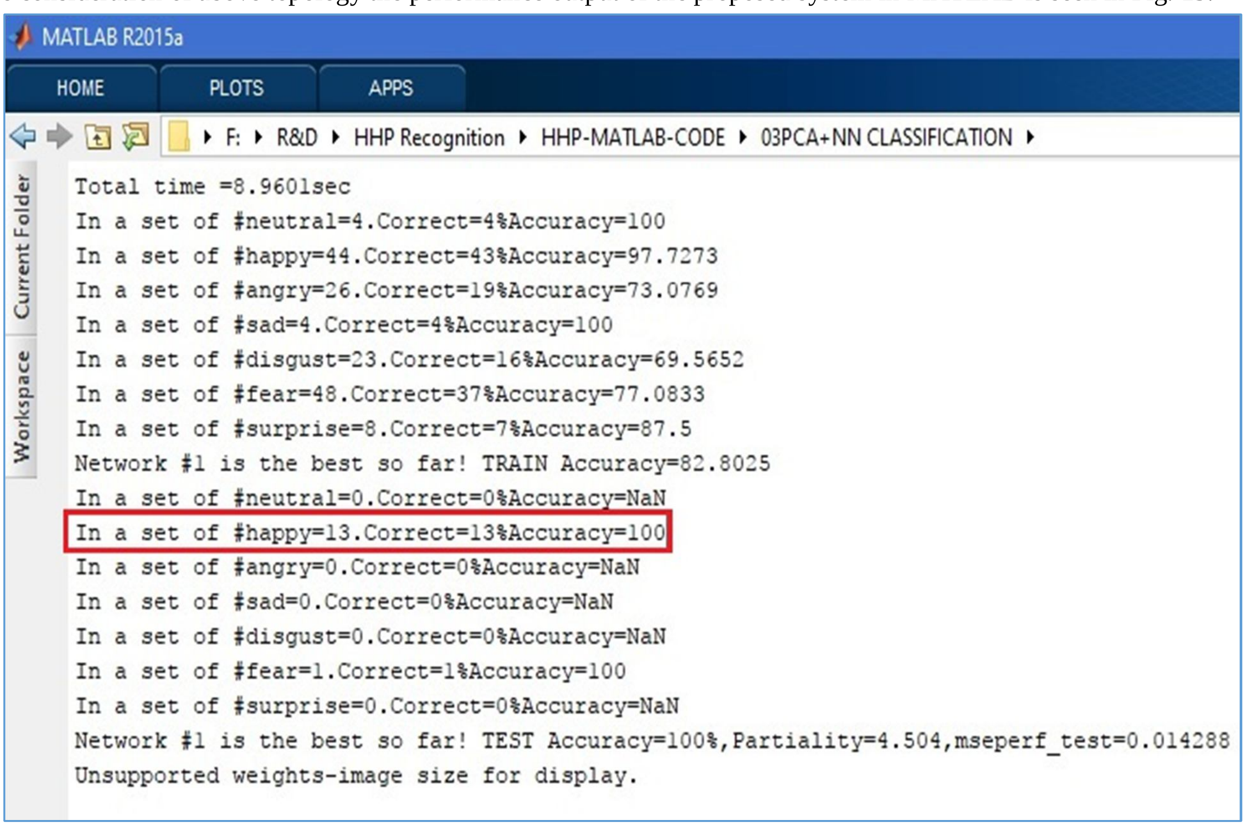


Fig. 13 MATLAB Output for the Emotion Happy of Subject 1

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

The MATLAB output for the subject 1 for a single emotion "Happy" is summarized in the following table. Which shows the accuracy level of training set classification is 82.82% as well as the accuracy level of testing set having only gesture images related to an emotion Happy of subject 1 is 100% with respective Mean Square Errors (MSEs).

TABLE III
MATLAB OUTPUT SUMMARY OF SUBJECT 1 FOR AN EMOTION HAPPY

Subject 1	Emotion	Neurons in hidden layer	Eigen Images Taken	No. of Iteration (Epoch)	Train Accuracy	Train MSE	Test Accurac y	Test MSE
	Happy	26	40	1600	82.82%	0.0034	100%	0.01428

In the same way the classification results of various hand gestures of subject 1 are classified over related individual emotions Anger, Sad, Disgust, Fear and Surprise including Neutral is tabulated as follows.

TABLE IV
MATLAB OUTPUT SUMMARY OF SUBJECT 1 FOR ALL BASIC SEVEN EMOTIONS

Subject 1	Emotion	Neurons in hidden layer	Eigen Images Taken	No. of Iteration (Epoch)	Train Accuracy	Train MSE	Test Accuracy	Test MSE
	Neutral	26	40	1600	80.2%	0.0039	100.0%	0.01426
	Happy	26	40	1600	82.8%	0.0034	100.0%	0.01428
	Anger	26	40	1600	83.2%	0.0032	82.0%	0.02629
	Sad	26	40	1600	84.3%	0.0029	81.0%	0.02814
	Disgust	26	40	1600	82.7%	0.0035	100.0%	0.01428
	Fear	26	40	1600	84.1%	0.0030	83.2%	0.02435
	Surprise	26	40	1600	92.4%	0.0024	97.3%	0.01632

As a sample of 10 individuals are taken into the experiment with their variety of finger, palm and arm said hand gestures related to basic six emotions including neutral position. The proposed system is simulated on all 10 subjects for all seven emotions separately and the obtained results can be seen graphically in the following figures.

Fig. 14 Experimental Sample of 10 Individuals

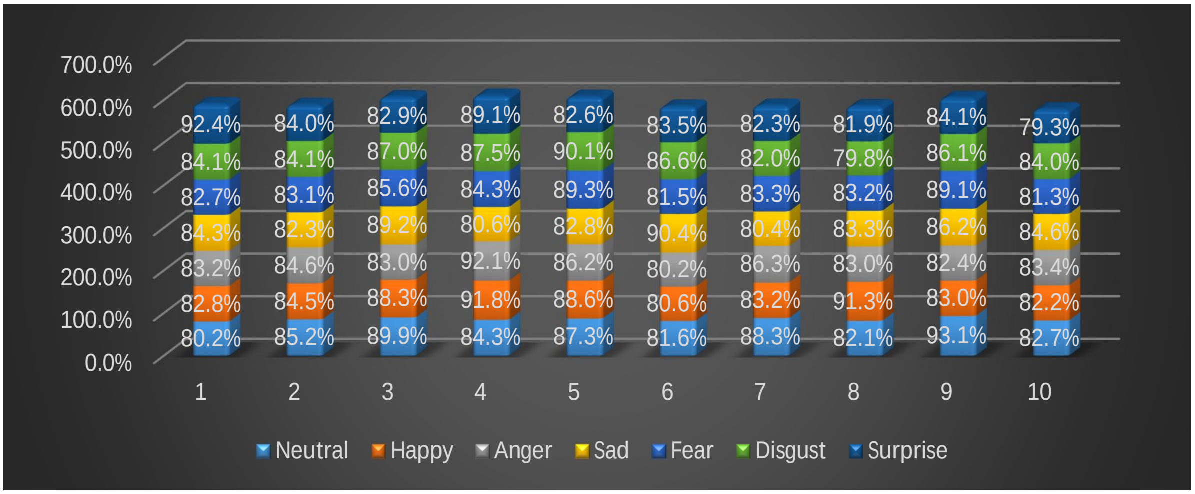


Fig. 15 Percentage Train Accuracy PCA + NN Hand Gesture

The experimental result shown in Fig. 15 indicates that the overall accuracy of training set of all the hand gesture images related to all seven emotions is 80-90%. Which means it suggests that some of the images of hand gesture for particular emotions are not classified into the category of respective emotion. Thus, there should more precise images be updated into the training set such that the level of accuracy can be improved. But, overall this result is still satisfactory as far as recognition is concern and has a scope of improvement.

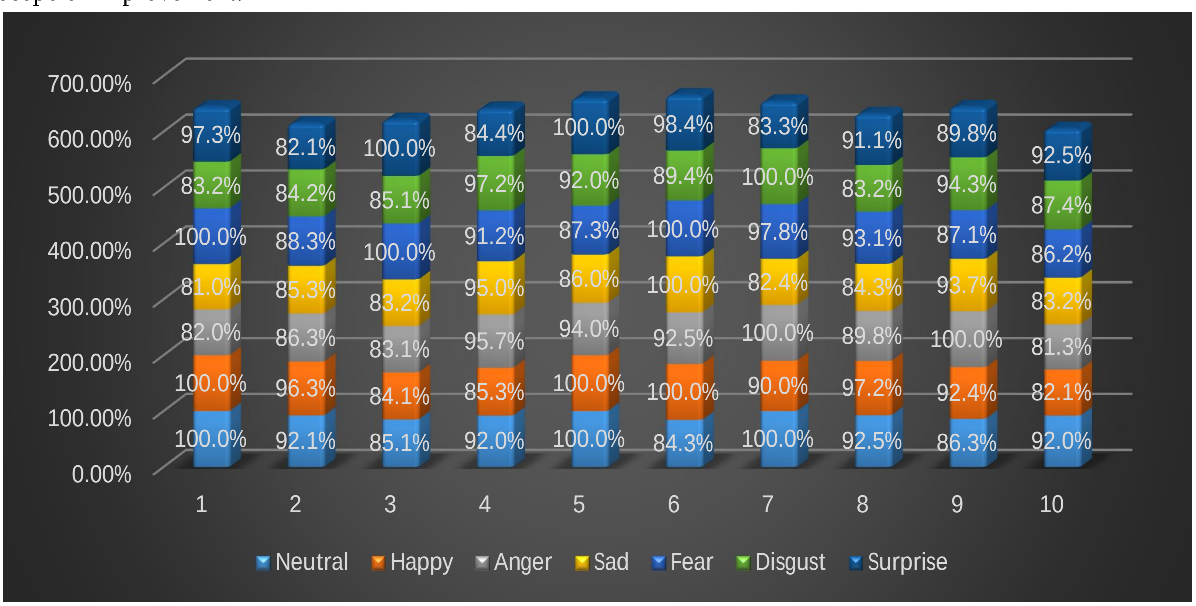


Fig. 16 Percentage Test Accuracy PCA + NN Hand Gesture

The outcomes of recognition of emotion through hand gestures shown in Fig. 16 indicates that the newly entered hand gesture test images of a particular emotion is classified over pre stored train image sets of the same emotions has the accuracy level between 90-100%. Which is very good, it is because of the experiment has been performed under well-defined environment. We should remember that it can be vary under real phenomena but that can also be managed by taking more precise images of hand gesture and more training of system such that the expected accuracy can be obtained.

IV. CONCLUSION

The accuracy level of recognition of an emotion, of an individual human being using PCAs of hand gestures classified against the basic seven emotions Neutral, Happy, Sad, Fear, Anger, Disgust and Surprise using Neural Network has been satisfactory. It is a kind of measurement of emotion which helps one to deal with counterpart. It also makes the communication process smoother, faster and more fruitful such that the decision making process becomes easier and the ultimate goal can be achieved.

REFERENCES

[1] Allan and Barbara, "The Definitive Book of Body Language, How to Read other's Thought by their Gesture", Pease International, ISBN 1-9208160-7-0, Australia, 2004.

[2] Konrad Schindler, Luc Van Gool, Beatrice de Gelder, "Recognizing Emotions Expressed by Body Pose: A Biologically Inspired Neural Model", Neural Networks Journal 21, pp. 1238–1246, 2008.

[3] David B. Givens, "The Nonverbal Dictionary of Gestures, Signs & Body Language Cues", Center for Nonverbal Studies Press, Spokane, Washington, 2002.

[4] Nele Dael, Marcello Mortillaro, and Klaus R. Scherer, "Emotion Expression in Body Action and Posture", American Psychological Association, DOI: 10.1037/a0025737, Vol. 12, No. 5, 1085–1101, 2012.

[5] Ekaterina P. Volkova, Betty J. Mohler, Trevor J. Dodds, Joachim Tesch, and Heinrich H. Bülthoff, "Emotion Categorization of Body Expressions in Narrative Scenarios", Frontiers in Psychology, Emotion Science, Volume 5, DOI: 10.3389, 2014.

[6] Gonzalez, R. C. and Woods, R. E., "Digital Image Processing", Prentice Hall, 3rd Ed. 2009

[7] Daw-Tunglin, Taiwan, "Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network" Journal of Information Science And Engineering 22, pp. 1033-1046, 2006.

[8] Jawad Nagi, Syed Khaleel Ahmed , "A MATLAB based Face Recognition System using Image Processing and Neural Networks", 4th International Colloquium on Signal Processing and its Applications, pp. 83-88, March 7-9, 2008.

[9] Rajasekaran, S. and VijayalakshmiPai, G.A.: "Neural Networks, Fuzzy Logic and Genetic Algorithms: Synthesis and Applications", Prentice Hall of India, 2003.

[10] https://in.mathworks.com