Journal of Basic and Applied Engineering Research
Print ISSN: 2350-0077; Online ISSN: 2350-0255; Volume 1, Number 6; October, 2014 pp. 61-66

http://www.krishisanskriti.org/jbaer.html

Human Face Emotion Recognition

Himani Sabharwal[1], Akash Tayal[2]

[1]*Student at Department of Electronics and Communication, IGDTUW, New Delhi, India*
himani.sabharwal987@gmail.com
[2]*Astt. Prof. , Department of Electrical Engineering, IGDTUW and USIT, New Delhi, India*
akashtayal@yahoo.com

Abstract: **SIFT (scale invariant feature transform) algorithm is famous and mostly used for object recognition which gives almost a high quality result. From past sometime it has been used on face to extract facial feature extraction and face recognition. In this paper I have applied fuzzy clustering as preprocessor, sift as facial feature extractor and svm as classifier. The keypoint features which are detected are used to recognize the emotions of human face via sm classifier. JAFFE (Japan Female Facial Expression) database. Results of this work are displayed in the form of percentage. Results of emotion recognition are diagnosed and displayed in such a way that higher the percentage of the emotion type in a single image in the gui prepared for human emotion recognition system; that will be the emotion diagnosed for the particular input image.**

Keywords: Emotion recognition, face recognition, fuzzy clustering, sift, svm.

1. INTRODUCTION

Human communication not only includes spoken languages or verbal words to express their feelings and emotions. There are other ways too from which one can assume the emotion of another person such as facial expression, hand gesture, tone of speech. To rectify these emotions computer is playing an important role of interaction between human and computer which developed human computer interaction system(HCI).

In all of the above emotion recognition ways face has always received a significant attention as human face is the richest source of human emotions[1]. Analysis of face expression benefits lawyers, police, security agents, entertaining industry. These people are interested in to know about the dishonesty, criminal behavior and attitude ; where as entertaining industry use these expressions to rectify and make their films or cartoons to entertain us.

Face emotion recognition system works in an order. For emotion recognition first thing is to diagnose the concerned area of image. Second, is to detect features of the face. Third, step is the comparison step where we compare the test and training images to get the type of emotion a face is showing.

2. RELATED WORK

Face has always been a fascinating part to work on. But the main problem every researcher faces is how to diagnose the key features of pattern or image under consideration. As there are many algorithms that can be used for feature extraction[6]. Based on what algorithm used decides the result of outcome. From [24], In 1981 Moravec worked on the development of image matching using corner detector. This work was further enhanced by Harris and Stephens in 1988. They made Moravec work more efficient for image variation and adge detection. In 1992 they also worked for motion feature tracking and 3d structure of motion.

In 1995 two researchers showed and developed same concept of matching corners by using correlation windows around each corner. In 1997 two more researchers found that this concept of feature extraction can be applied to general image recognition. In 1999 David lowe extended this work with a new approach for stable keypoints. This is not a full history of the SIFT algorithm. Some of the earlier work also identified the peaks and ridges in scale space and linked it into tree structure[24]. After the feature extraction the classifier is another important module for facial emotion recognition system[7].

Researchers are nowadays trying to recognising the emotion in multiple pattern. Mainly the emotions being worked mainly on six type of them: anger, disgust, happy, fear, sad and surprise. But there lot more emotions combined for the same emotion appearing at a particular emotion. Svm is the best option to use as a classifier for multiple emotion identification. It is based on statistical learning theory and is being used mostly by researchers nowadays.

3. THEORY

Before we go to know about the experiment performed and the result attained we have to go through the theory of the algorithms used by us. So the discussion about the algorithms consists of three algorithms they are as follows:

A. Fuzzy clustering(FCM)

Fuzzy logic has become very famous in past few years for pattern recognition. Pattern recognition can be of any type like object, human organs, face. Clustering basically is defined as a technique for image segmentation. FCM is fuzzy c-mean algorithm which is one of the methods of fuzzy clustering.

FCM allows a particular point to partially belong to more than one cluster at the same time. Fcm is used to modify an objective function($J_m(p,v)$) which evaluates the compactness and distance between the clusters formed. The objective function of FCM is given as :

$$J_m(p,v)=\sum_{i=1}^{k}\sum_{xk\epsilon X}\left(\mu ij(xk)\right)^m \left|\left|xk - vi\right|\right|^2 \quad \ldots\ldots \quad (1)$$

Where p is fuzzy partition, X is dataset, C is no. of clusters, m is fuzzy weighting exponent for μ_{ji} which controls the fuzziness of the resulting clusters.

B. SIFT Feature Extraction

SIFT is elaborated as Scale Invariant Feature Transform. SIFT has been successfully applied to image matching to feature extraction and face recognition. It extracts features by detecting the keypoints of the image. As described in [6] SIFT algo consists of 4 basic steps. These are as follows :-

Scale Space Extrema Detection – This step is used to detect keypoints of image. These keypoints are basically the interest points which are important for us. To detect keypoints image I(x,y) is convolved with Gaussian blurs G(x,y,σ) at different scale to get Gaussian images. This function is given as [24] :

$$L(x,y,\sigma) = G(x,y,\sigma)*I(x,y) \quad \ldots\ldots \quad (2)$$

These zoomed images of original image by a factor of σ, kσ…..etc are subtracted from each other to get the keypoints of the image. This difference of zoomed images by series of factor is called difference of gaussian images(DOG). These keypoints which are extracted from series of image are maxima and minima that occurs at multiple scales. Thus it is an feature improving procedure where the subtraction of one blurred image from another more blurred image occurs. The formula used on above concept is [24]:

$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \quad \ldots\ldots \quad (3)$$

Keypoint Localization – keypoints obtained from scale space extrema detection consists of keypoints – stable and unstable. This step detect unstable keypoints and eliminates to obtain stable and reliable keypoint features. Unstable keypoints are eliminated due to low contrast, poor edges and below threshold value[5-7].

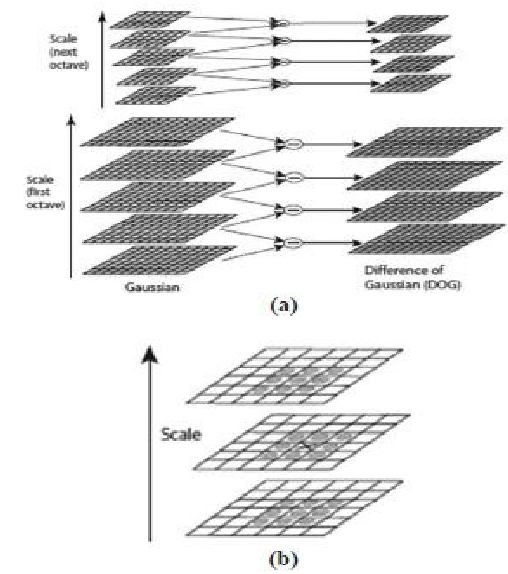


Fig. 1(a) representation of series of difference of Gaussian. 1(b) neighborhood pixel values are compared to get keypoints

Orientation Assignment- According to this step, it gives more than one orientations based on local image gradient directions. These directions are calculated[24] as

$$\boldsymbol{m(x,y)} = \sqrt{\left[\left(\boldsymbol{L(x+1,y)} - \boldsymbol{L(x-1,y)}\right)^2 + \left(\boldsymbol{L(x,y+1)} - \boldsymbol{L(x,y-1)}\right)^2\right]} \quad \ldots\ldots \quad (4)$$

$$\theta(x,y) = \tan^{-1}\frac{[L(x,y+1)-L(x,y-1)]}{[L(x+1,y)-L(x-1,y)]} \quad \ldots\ldots. \quad (5)$$

Steps involved in calculating them are[6] as follows:

- A neighbourhood M is choosen by the center of the keypoint.
- Now the directions of keypoints in M is calculated by using equation (5).
- Direction distribution is calculated.
- Statistical histogram is also drawn.
- With the help of statistical histogram direction of keypoint is diagnosed.
- The direction of keypoint is that quantity that consists of largest histogram bar on it's division.

- Keypoint Descriptor – Above 3 steps gives the scale, position and direction respectively. These steps can only ensure the geometric compatibility. To ensure illumination, view perspective related problems this ste is implemented. To implement this a 16x16 pixel neighbourhood field of M is selected. This field is divided into 16 subfields of section 4x4. Using (4) and (5) m(x,y) and $\theta(x, y)$ is calculatedfor every subfield. Using (5) 8 direction descriptors are found with the help of histogram plots. Feature descriptor keypoints are deduced by connecting all subfields and calculating their directions.the total of the direction description is 16 and length of the feature description is 16x8=128 [6]. Feature descriptors are normalized to eliminate illumination problem.

C. Support Vector Machine(SVM)

Support vector machines are based on studies of statistical learning theory, pioneered by vapnik[7]. SVM is used as a classifier, which is basically used to classify the input data when compared to database clarifies the category of the input data. But the results derived are always depends upon the complexity of problem statement. This algorithm is based on mathematical concept. The function f(x) determines the margin or hyperplane[8]. Hyperplane separates two or more classes of dataset and datapoints of an image. Separated data points helps in diagnosing the category of the input data.

SVM is calculated for 2 types of dataset- linear and non-linear. For linear svm classifier function f(x) is given as [9]:

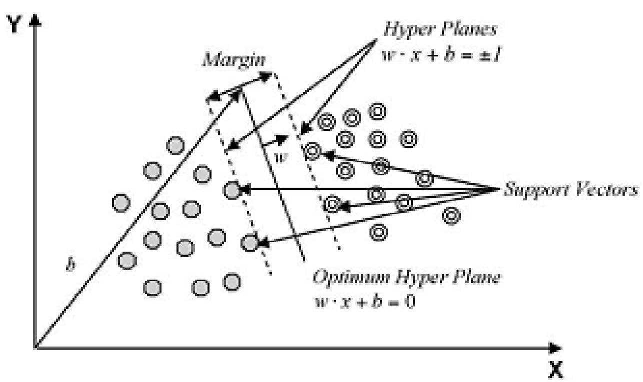


Fig. 2 Hyperplane – margin representation

$f(x;w,b)=<w,x>+b$ …… (6)

here f(x) is a classifier, (x_i,y_i) are set of points where i=1, …..,l; x_i □ R^d and $y_i \epsilon [-1, +1]$, <,> denotes inner product of 2 vectors, w and b are parameters of function; w is the normal vector to the hyperplane and x is the parameter that determines the offset of the hyperplane from the origin along the normal vector w.

$x=\frac{b}{|w|}$ …… (7)

data points induced to get f(x) for both the extreme boundaries is given as:

$<w,x_i> + b_0 \geq 1$; for $y_i = 1$ …… (8)

$<w,x_i> + b_0 \leq -1$;for $y_i = -1$ …… (9)

Equation (9) can be modified as
$Y_i[<w,x_i> + b_0] - 1 \geq 0$ …… (10)

To find maximum margin hyperplane the following optimization problem should be solved:

$\text{Min}_{w,b} \frac{<w,w>}{2} = \min \frac{||w||^2}{2}$ …… (11)

With reference to equation:
$Y_i[<w,x_i> + b] \geq 1$ …… (12)

It can also be solved by using lagrange function,α_i

$$L(w, b, \alpha) = \frac{1}{2} < w, w > - \sum \alpha_i [y_i(< w, x_i > +b) - 1]$$

....(13)

To find appropriate α_i, Lis minimized with respect to w and b by taking their derivatives with respect to w and b respectively:
$w = \sum_{i=1}^{l} y_i x_i \alpha_i$ …… (14)

$0 = \sum_{i=1}^{l} y_i \alpha_i$ …… (15)

Putting equation (14 , 15) into (13) we get Maximum

$w(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} y_i \, y_j \alpha_i \alpha_j < x_i, x_j >$ 16)

for equation (15), lagrange function for conditions from equation (10 to 16) is calculated as:

$\alpha_i [y_i(< w, x_i > +b_0) - 1] = 0$ …… (17)

L is maximized with respect to non-negative α_i ,

$\text{Max} dl = -\frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j < x_i, x_j >$ (18)

$f(x) = \sum_{i=1}^{l} y_i \alpha_i < x_i, x > + b$ (19)

Those data points that lie on the hyperplane are called support vectors. Till now we have discussed about linear svm classifier, but if data is non-linear and inseparable then svm is used with the introduction of a slack variable to the same equations of linear svm classifier. The idea behind it is that those points which have got placed on the wrong side of the hyperplane are placed back by introducing the slack variable. Slack variable is used to control the value of distance by which data points are wrongly place across hyperplane.

4. HOW EMOTION RECOGNITION SYSTEM WORKS

Figure 3 below describes the human emotion recognition module

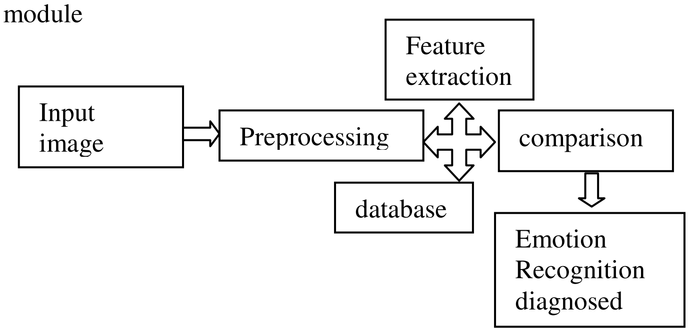


Fig. 3. emotion recognition module

a) Database: It contains images of JAFFE female database for 5 emotions – anger, disgust, happy, neutral and sad. Every time when an input is given, it is compared to this database and gives an output as emotion detected.

b) Preprocessing: The main aim of this step is to have focus on interested region of image. For this i have used fuzzy clustering via FCM that created 5 clusters of face to have focus only at concerned area of image here it is face in this case.

c) Feature Extraction: This step is the most essential part of emotion recognition module. For feature extraction SIFT is used. It extracts the stable keypoints as feature descriptor of image. Feature of every image in certified database is extracted and saved as database. This database is used to compare for every input image.

d) Comparison: Emotion for every input image is analysed and declared by comparing it with certified database which passes through a classifier as SVM in this case.

2) Experiment:

In this paper, to display the implementation of my work a GUI is prepared. This GUI consists of :

a) pop up buttons – This GUI system consists of 4 buttons – input image, clustering, sift feature, results. Every popup button when clicked display their final output display at axes. But final button when clicked i.e. results display the percentage of each type of 5 emotions. The type of emotion which has highest percentage is the real emotion diagnosed for the input image on comparision with the database.

b) Axes: This region is the area where all the outputs of images are displayed. It is placed beside the popup buttons.

Static and Edit text box: Static text box displays the name of 5 types of emotions whereas edit text box is displayed infront of every emotion which displays the percentage of every emotion type in a single input image.

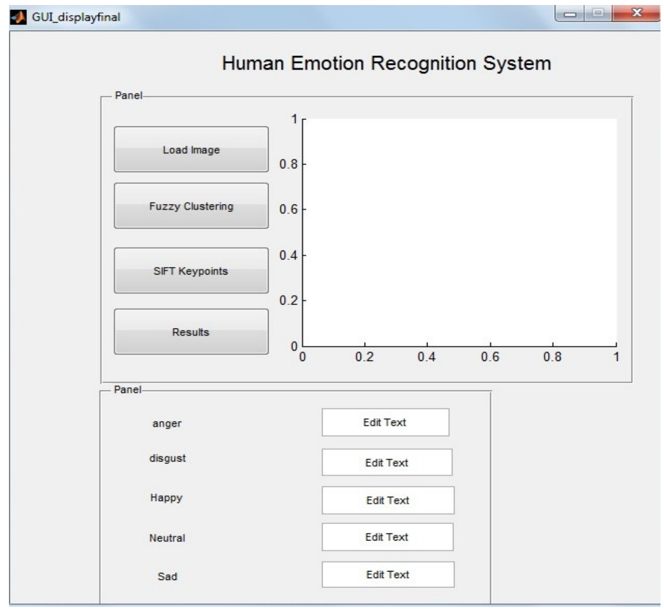
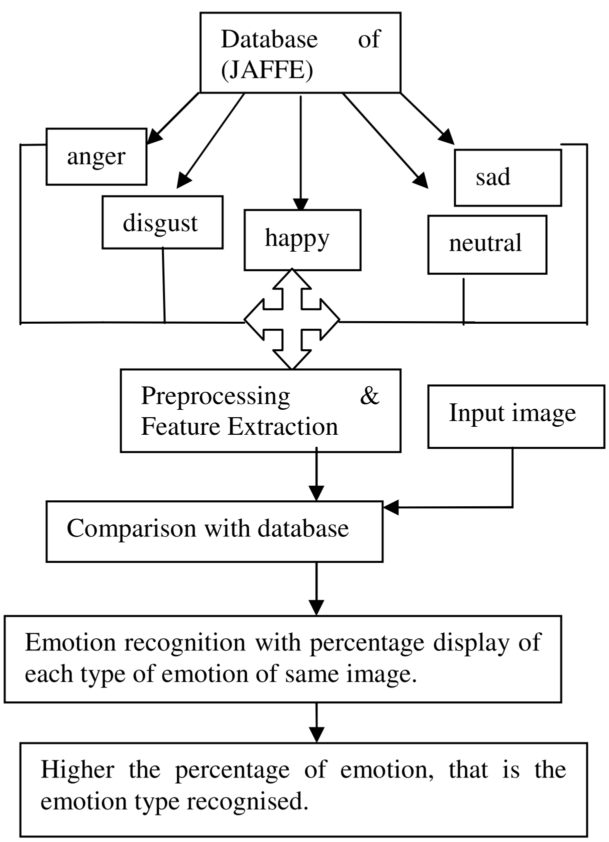


Fig. 4 Human emotion recognition system

3) Proposed Algorithm

The explanation of how my system of emotion recognition of human face works can be examined here. The algorithm involved is:



Algorithm :

1) Train data is loaded .
2) Now for each feature extracted from jaffe images in jaffe database were provided path.
3) For example for a=1 path of jaffe database as excel sheet was provided , then read and their mean were calculated and again were stored as in excel sheet feature extraction.
4) This was done for every folder of emotions.
5) Mean was stored as data(i).
6) These data were stored in final_data.
7) Then 5 classes were created and was concatenated.
8) Now using multisvm
9) [itrfin]=multisvm(final_data,tot_train,final_data)
10) Where final_data is the data collected (mean calculated)
11) Tot_data is the class created.
12) Along with svm, fuzzy and sift is used.
13) After this difference between class data and transpose of svm values calculated and stored in that variable. And then using the values they are recognized as 5 emotions – anger , disgust , happy , neutral & sad.
14) Test data : In test data folder there are 10 images of multiple emotions in it .
15) Jaffe data (trained data) :

    a) In this data folder there are 5 sub folders with naming of 5 emotions – anger , disgust , happy , neutral & sad.

    b) These 5 folders contain 5 images each of 5 emotion i.e. in total 0f 25 images in trained data each classified with their emotions .

5. RESULT AND CONCLUSION

In this experiment JAFFE database is used. This database encloses 5 types of emotions. These emotions are anger, disgust, happy, neutral and sad. Emotion is categorized by making its separate folder. Each emotion folder contains 5 images. Thus in total we have taken 25 images for dataset to create. Out of these 25 certified images 10 images are used for test data to find accuracy of our work. Implementation of this system gives performance of about 98% successful recognition of emotion. The result for an image is as follows:

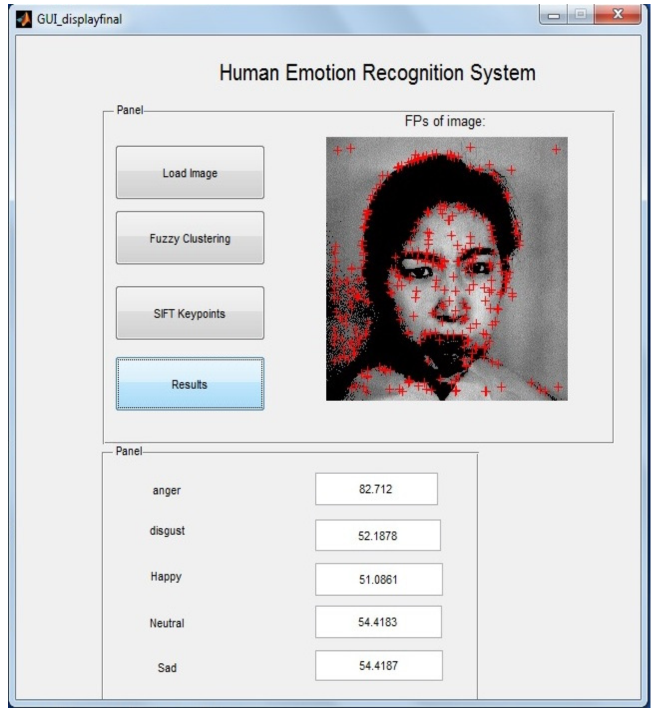


Fig. 5. Result of GUI system display

The result of this emotion recognition coding where test data images are being examined is almost giving the correct emotion displayed by input image. To enhance my result I have converted this work into GUI display system which displays percentage of each emotion in a single input image. The emotion with higher percentage is the main emotion of the image being analysed.

6. ACKNOWLEDGMENT

This work was supported by Mr. Akash Tayal under his guidance at IGDTUW. I would like to thank him for suggestions in improvement to both the content and work of this paper.

REFERENCES

[1] Assia Khanam, M. Zubair Shafiq and M. Usman Akram, “Fuzzy based facial expression recognition” , 2008 IEEE Congress on Image and Signal Processing.

[2] M. Karthigayan, M. Rizon, Sazali Yaacob, R. Nagarajan, M. Sugisaka, M.Rozailan Mamat and Hazry Desa, “Fuzzy clustering for genetic algorithm based optimized ellipse data in classifying face emotion,” in ICROS International conference on

[3] control, automation and systems 2007, oct. 17-20 in coex, seoul, korea.

[4] Qing Zhang and Minho Lee,“Fuzzy-GIST for emotion recognition in natural scene images,” in 2009 IEEE 8th international conference on development and learning.

[5] Ran Zhou, Jie Wu, Qing He, Chao Hu, and Zhuliang Yu, “Approach of human face recognition based on sift feature

extraction and 3d rotation model," in IEEE 2011 at international conference on information and automation Shenzhen,china in june.

[6] Cong Geng and Xudong Jiang, "SIFT Feature for Face Recognition," in 2009 at IEEE.

[7] Han Yanbin, Yin Jianqin and Li Jimping, "Human Face Feature Extraction and Recognition Base on SIFT," in 2008 international symposium on computer science and computational technology, IEEE.

[8] Hui Zhao, Zhiliang Wang and Jihui Men, "Facial Complex Expression Recognition Based on Fuzzy Kernel Clustering and Support Vector Machines," third International Conference on Natural Computation (ICNC 2007) 0-7695-2875-9/07 , IEEE.

[9] Dmitriy Fradkin and Ilya Muchnik, "Support vector machines for classification," at DIMACS series in discrete mathematics and theoretical computer science, 2000 mathematics subject classification,62H30.

[10] Spiros V. Ioannou, Amaryllis T. Raouzaiou, Vasilis A. Tzouvaras, Theofilos P. Mailis, Kostas C. Karpouzis and Stefanos D. Kollias, "Emotion recognition through facial expression analysis based on a neurofuzzy network," in 2005 special issue, S.V. Ioannou et al. / Neural Networks 18 (2005) 423–435.

[11] Deepthi.S, Archana.G.S and Dr. Jagathy.V.P, " Facial expression recognition using artificial neural networks," in IOSR journal of computer engineering, ISSN: 2278-0661, ISBN: 2278-8727Volume 8, Issue 4 (Jan. - Feb. 2013), PP 01-06.

[12] Shail Kumari Shah, " A survey of facial expression recognition methods," in IOSR journal of engineering (iosrjen), ISSN (e): 2250-3021, ISSN (p): 2278-8719 Vol. 04, Issue 04 (April. 2014), ||V3|| PP 01-05.

[13] Neeraj Singla and Sugandha Sharma, " Advanced survey on face recognition techniques in image processing," volume 3 issue 1, January 2014 ijarcet.

[14] Faizan Ahmad, Aaiman Najam and Zeeshan Ahmed, " Image-based face detection and recognition: state of the art," IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 6, No 1, November 2012.

[15] Dilbag Singh, " Human emotion recognition system," in august 2012 in MECS, DOI 10.5815/ijigsp.2012.08.07.

[16] Amav Bhavsar and Hima M. Patel, "Facial expression recognition using neural classifier and fuzzy mapping," IEEE Indicon 2005 Conference, Chennai, India.

[17] Ligang Zhang , Dian Tjondronegoro and Vinod Chandran, "Facial expression recognition experiments with data from television broadcasts and the World WideWeb," Image and Vision Computing 32 (2014) 107–119.

[18] Aruna Chakraborty, Amit Konar, Uday Kumar Chakraborty, and Amita Chatterjee, "Emotion recognition from facial expressions and its control using fuzzy logic," IEEE Transaction on systems, man, and cybernetics — part A: systems and humans, VOL. 39, NO. 4, JULY 2009.

[19] Anisha Halder, Rajshree Mandal, Amit Konar1, Aruna Chakraborty and R. Janarthana, "Emotion recognition from facial expression using general type-2 fuzzy set," 2011 International Conference on Recent Trends in Information Systems.

[20] Wenyu Chen, Yanli Zhao , Wenzhi Xie and Nan Sang, "An Improved SIFT Algorithm for Image Feature-Matching,"

[21] David G. Lowe, "Object Recognition from Local Scale-Invariant Features," Proc. of the International Conference on Computer Vision, Corfu (Sept. 1999).

[22] Chai Tong Yuen , Woo San San , Mohamed Rizon and Tan Ching Seong, "Classification of human emotions from eeg signals using statistical features and neural network," International Journal of Integrated Engineering (Issue on Electrical and Electronic Engineering).

[23] Chul Min Lee and Shrikanth Narayanan, "Emotion recognition using a data-driven fuzzy inference system.

[24] David G. Lowe, "Local feature view clustering for 3d object recognition," Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii (December 2001).

[25] David G. Lowe, "Distinctive image features from scale-invariant keypoints," in international journal of computer vision, January 5, 2004.

[26] I. Guyon, J. Weston, S.Barnhill and V.Vapnik, "Gene selection for cancer classification using support vector machines," machine learning, 46(1-3):389-422, 2002.

[27] K.Crammer and Y. Singer, "On the learnability and design of output codes for multiclass problems," in computational learning theory, 2000.

[28] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines and other kernelbased learning methods," Cambridge university press.

[29] J. Weston and C. Watkins, "Multi-class support vector machines," department of computer science, university of London, 1998, technical report csd-tr-98-04.