Sensor and Feature Level Fusion of Thermal Image and ECG Signals in Recognizing Human Emotions

C. M. Naveen Kumar, G. Shivakumar

1

Abstract— Recent studies on recognition of various emotion labels concentrated on speech signals, text, visual images and anatomical variables. The proposed system combines the features of ECG which are extracted using empirical mode decomposition and features of thermal images which are extracted from Gray Level Co-occurrence Matrix (GLCM) viz energy, contrast, homogeneity and correlation. ECG is acquired from AD8232 module and thermal images from FLUKE TiS20. Data of ECG and thermal images are acquired simultaneously from a subject and database consists of data from 40 subjects in age group of 20 years to 40 years from Hassan, Karnataka, India. Here different labels of emotions have been classified based on K-nearest neighbor decision rule. This system yielded highest accuracy for disgust and lowest for anger using ECG and highest accuracy for disgust and surprise and least for sad.

Keywords—human emotions, electrocardiogram, thermal image, empirical mode decomposition

I. INTRODUCTION

Emotions are integral part of human's social and daily life. In today's modern world emotional intelligence is a buzz word which is the ability to consciously aware of one's emotional expression and to control emotions in order to maintain interpersonal relationships. From theory of psychology, primary human emotions are classified into six types they are: happy, surprise, fear, sad, anger and disgust. According to Paul Ekman, Human beings normally intermittently hide their states of emotions in order to avoid the situation which can cause embarrassment. People control their instant emotional expressions according to their gender, cultural associations and background. Normally fake smile is exhibited in place of anger or sadness since they are not socially encouraged. Anatomical variable analysis may be used to identify intentionally hidden affective states [1].

Human Emotions and the literature related to it have its foot prints in many disciplines like computer science, psychology, and human physiology and can be understood by the detailed investigation. Recent developments in technology allowing researchers to analyse human brain and circuitry associated with activity of emotions which can help in detailed understanding of human brain's structure and working related to emotions and also in recognition of emotions. Current developments in signal processing and analysis permitted us to analyse physical signals which is correlated with emotions and thus emotion recognition interfaces can be built [2].

Identification of Emotions is a complicated problem and till now scientists and researchers in the area of psychology and theory of emotions have not came up with clear solution to this. Emotions can be identified using facial expression, voice, textual data and also by statistical analysis of human anatomical variables. Human's activity is related bioelectricity that reveals excitability, which is a vital activity of life and can be seen in all stages of life. Change in the physiological activity results in electrical signals which are deeply correlated to human emotions. Currently, human emotion analysis from anatomical variables focuses on statistical characteristics in both frequency and time domain [5].

Thermal Imaging can be contact less method to evaluate human emotions since radiations from face and its topographical distribution of temperature is correlated to various affective states.

II. RELATED WORK

Social Robots requires an ability to recognize human emotions and is considered to be a key factor. Earlier researches focused on modalities that yielded lower recognition rates and faced many issues while applying the same into the robots. Decision level fusion method is proposed where facial expression is recognized through Convolutional Neural Networks and emotional content in the speech is detected from speech analytics engine. Here features that gives best rate of recognition is combined and used as input to the classification stage for ANN as well as k-NN[3].

The essential features of Electrocardiogram which are necessary in diagnosis of heart and cardiovascular activity can be revealed through the signal processing techniques which may be in Time and frequency domains. Fast track and well-organized treatment of cardiovascular disease can be done through classification of ECG. Firstly, Empirical Mode Decomposition (EMD) is used in decomposing acquired ECG data that yields Intrinsic Mode Functions (IMF) of higher order. Modified ECG is generated using combination of first three IMFs.

Revised Manuscript Received on December 12, 2019.

***** Correspondence Author

C. M. Naveen Kumar*, Dept. of E & I Engg, Malnad College of Engineering, Hassan affiliated to Visvesvaraya Technological University, Belagavi, India. Email: cmnk.it@gmail.com

G. Shivakumar, Dept. of E & I Engg, Malnad College of Engineering, Hassan affiliated to Visvesvaraya Technological University, Belagavi, India. Email: gs@mcehassan.ac.in

EMD has better noise suppression characteristics and also produces wide range of information. One dimensional CNN is designed that identifies and learns the features of modified ECG signal better than raw ECG and is used in classification stage which uses softmax regressor [4].

Emotion analysis of children is done from thermal infrared imaging utilizing emissivity variation in face. Study considered both valence and arousal dimensions of emotions. Thermal variations in regions like nose, cheek and chin is observed for various emotion labels and thermal depression in nose is found have association with valence. The variations in facial to expressions cause errors in Region of Interest (ROI) positioning and muscle deformation. This can cause thermal asymmetry in face. Manual Tracking and positioning of ROI is done to make sure the perfect positioning of ROI [6].

Human emotions that can be evaluated in a contactless manner using non-invasive technique is implemented in [8] that uses thermal signatures from face image from the front. A thermal imager which is a seamless device is utilized in order to detect various emotion labels where it registers heat signatures radiated from subject’s face. Texture analysis based on a unique algorithm which extracts feature from GLCM method which detects region where flows dominantly. In this region, first order statistics is computed based on the cues and this is combined with cues of higher order statistics to build an efficient feature vector that can expose important and dominant patterns for different labels of emotion [8].

III. METHODOLOGY

The block diagram in the figure 3.1 shows the proposed methodology.

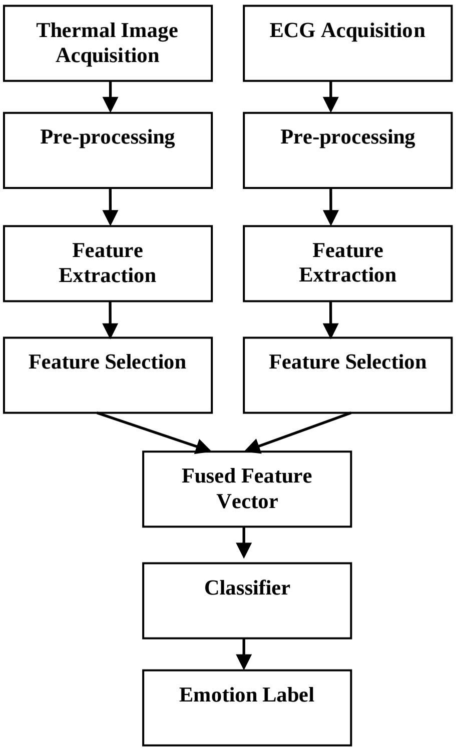

Fig. 3.1 Proposed system block diagram

A. DATA Acquisition

In order to recognize the emotion experienced by the subject, data samples must be acquired from them. Hence to evoke particular emotion in a subject, videos for each emotion were played. These videos were selected by conducting a survey for different age groups and collected their feedback. The duration of these videos is 180 seconds.

For real time data acquisition another survey was conducted. When the video is played, the subject starts to experience and expresses certain emotions. This causes change in their physiological signals and also temperature of the face. These changes are measured using ECG module. The setup made for data acquisition is shown in figure 3.2

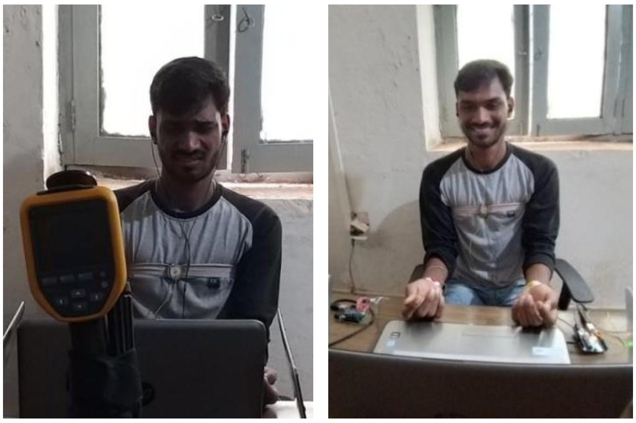

Fig. 3.2 Set-up for Data Acquisition from subjects

Thermal Image Acquisition: Thermal camera is placed at a distance of 1m from the subject. The images are taken for every 3seconds in an Auto capture feature mode.

ECG Data *Acquisition*: This signal is acquired using AD8232 module. Electrodes of this module are placed on the subject with respect to Einthoven’s triangle. The module is connected to arduino board and then it is interfaced with MATLAB 2018a. The program for acquiring ECG signal is written in MATLAB 2018a. After acquiring, data will be stored in Notepad.

B. Preprocessing

ECG signal: In handheld or wearable ECG acquisition system, signal degradation occurs due to drift in the baseline and artefacts related to slight displacement of ECG electrodes due to muscular activity. The drift in baseline can be eliminated using a high pass filter. In case if artefacts related to muscular activity dominates the ECG then the frequency components of these artefacts may get into desired portion of the acquired ECG signal. Median filter gives more robust value and will not be affected significantly by very unrepresentative sample in the neighborhood. It also reveals the underlying baseline drift [9]. The below figure 3.3a and 3.3b shows signal before and after pre-processing.

Retrieval Number: B10541292S19/2019©BEIESP
DOI: 10.35940/ijitee.B1054.1292S19

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

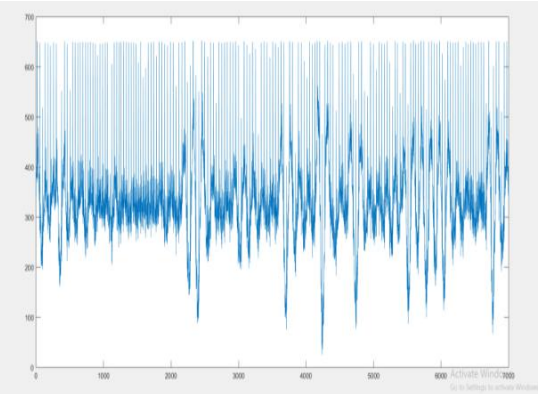

Fig. 3.3 a: Acquired ECG signal before preprocessing

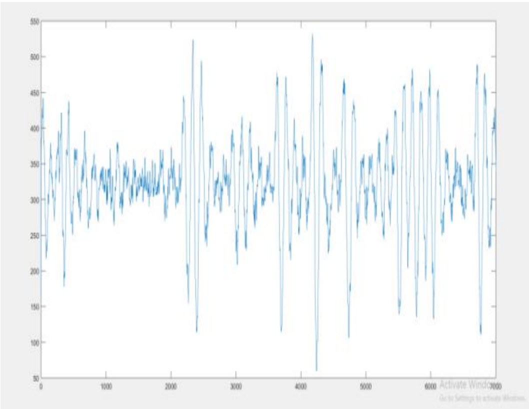

Fig. 3.3 b: ECG signal after preprocessing

<u>*Thermal images*</u>: Image is smoothened using a low pass Gaussian filter. A ROI is that part of an image that can be used for further processing. Face of the subject is the Region of Interest here which is cropped manually from each thermal image [8].

C. Feature Extraction & Selection

<u>*ECG signal*</u>: After applying Empirical Mode Decomposition (EMD) on ECG signal, three Intrinsic Mode Function (IMF) were considered viz IMF1, IMF2, IMF3. Due to bandwidth limitation of ECG signal IMFs were limited to three which are shown in figure 3.4 a to c.

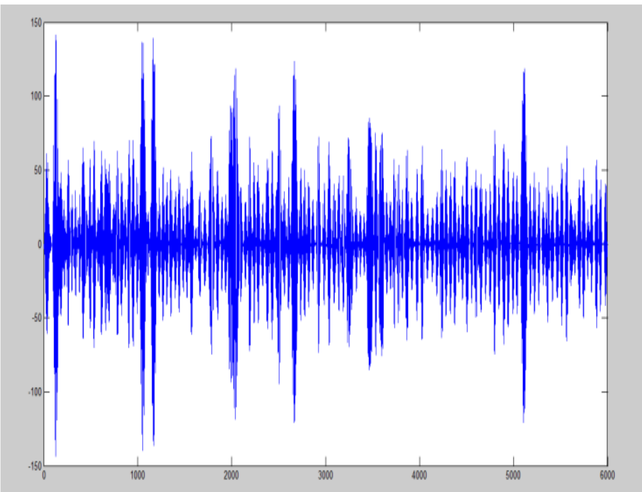

Figure 3.4 a: IMF1

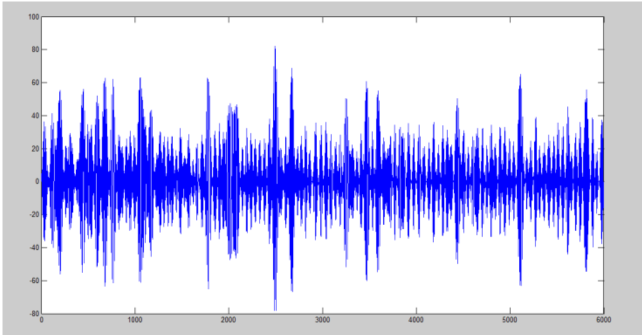

Figure 3.4 b: IMF2

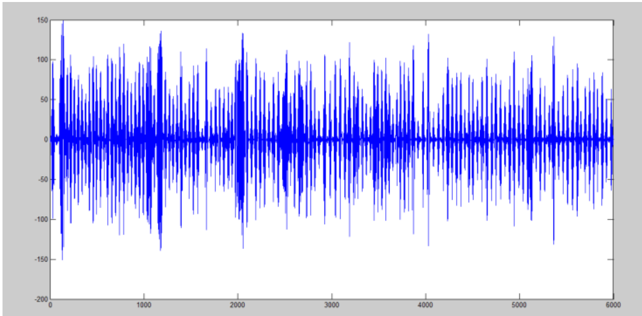

Figure 3.4 c: IMF3

From these IMFs three features were selected, they are: Mean, RMS and Maximum value. Totally 9 features were selected; 3 features from each IMF.

<u>*Thermal images*</u>: GLCM is used for performing statistical analysis to evaluate and compute textures in order to take the spatial connection of the pixels in to the consideration. It can be viewed as gray level histogram in two dimensions. Here pixel distance and direction is used to compute occurrence probability between a pair of pixels [8]. The functions based on GLCM categorize the image surface by considering pair of pixels of specific relationship between them in spatial domain and also specified values which is there in GLCM image forms [7]. For each thermal image, 16 matrices of order 4*4 are considered. 4 features are extracted from each matrix. These four features are co-relation, contrast, homogeneity and energy.

<u>Contrast</u>: It is used to compute the local variations in the GLCM. The value returned corresponds to contrast in intensity between a pixel and its neighbour over an entire input image. Here N is number of quantization levels. Here Element of normalized symmetrical GLCM is denoted by p(i,j).

$$\sum_{i,j=0}^{N-1} \quad (i-j)^2 p(i,j)$$
(1)

<u>Correlation</u>: In specified pair of pixels, it computes the joint probability of occurrence . The value returned corresponds to how a pixel is correlated with its neighbour over an entire input image. μ is the mean and σ^2 is the variance of all related pixels of reference that constitutes GLCM.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$\sum_{i,j=0}^{N-1} \left(\frac{(i-\mu)(j-\mu)}{\sigma^2}\right) \quad (2)$$

$$\mu = \sum_{i,j=0}^{N-1} i.p(i,j)$$
(3)

$$\sigma^2 = \sum_{i,j=0}^{N-1} (i-\mu)^2 p(i,j)$$
(4)

Range = [-1 1]

Energy: It is a measure of uniformity and computes the square of all elements in the GLCM and return its sum in the range = [0 1].

$$\sum_{i,j=0}^{N-1} p(i,j)^2$$
(5)

Homogeneity: It evaluates how closely the elements are distributed in GLCM with respect to GLCM diagonal. It returns a value in the range [0 1].

$$\sum_{i,j=0}^{N-1} \frac{p(i,j)}{1+(i-j)^2}$$
(6)

Data Fusion: Features from thermal ir image and ECG is obtained separately and fused in this stage. The method used to fuse these features at decision stage. The total number of feature extracted from thermal images are 64 and from ECG are 9. Totally 73 features are obtained after fusion.

D. Classification

The fused feature data which is pertinent to actual affective states is used for classification. Firstly the classifier is trained with the fused feature data then trained classifier recognizes the genuine state of emotion. In the proposed emotion recognizer, algorithm used for classification is K-Nearest Neighbor (KNN) decision rule which is a non-parametric technique whose output is a class membership. Emotional state is categorized based on the majority of vote by the neighbors to a particular class which is utmost conjoint amongst the k-nearest neighbor. This algorithm is used because of its robustness to the noisy input data and also effective for large input data. In this stage the fused features converges into 6 classes where each class represents a particular emotion.

IV. RESULTS & DISCUSSIONS

Recognition rate for six emotion labels is computed for ECG and for combined ECG and Thermal Image using database which is built and is tabulated in below table I.

Table-I: Recognition Rate for various Affective states

Parameters		ECG	ECG + Thermal Image
Emotion Labels	**Happy**	6	6
	Sad	6	6
	Disgust	6	6
	Anger	6	6
	Fear	6	6
	Surprise	6	6

Figure 4.1 to 4.6 shows the screen shots of GUI for various emotion labels.

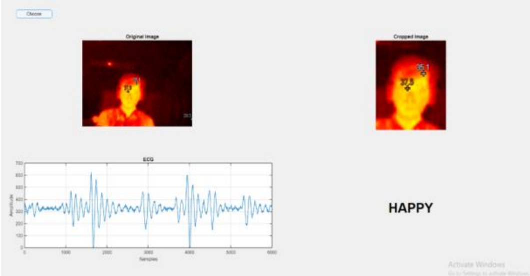

Fig. 4.1: Screen shot of GUI for HAPPY

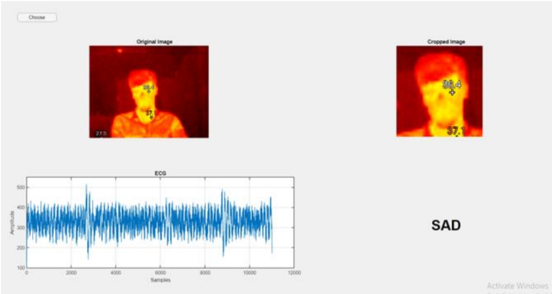

Fig. 4.2: Screen shot of GUI for SAD

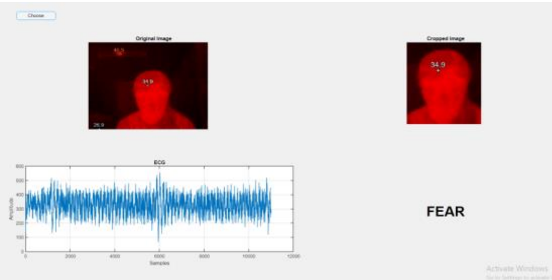

Fig. 4.3: Screen shot of GUI for FEAR

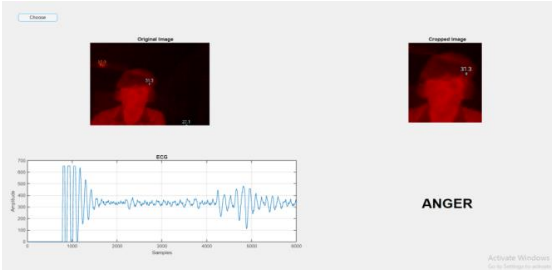

Fig. 4.4: Screen shot of GUI for ANGER

Retrieval Number: B10541292S19/2019©BEIESP
DOI: 10.35940/ijitee.B1054.1292S19

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

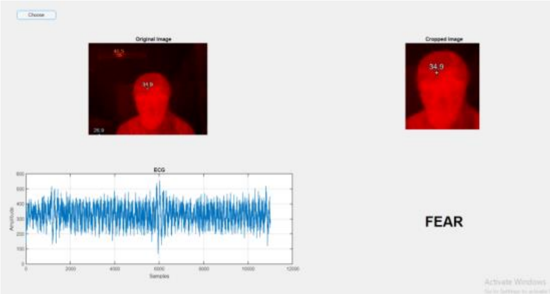

Fig. 4.5: Screen shot of GUI for FEAR

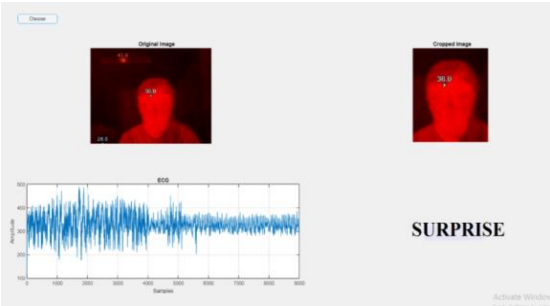

Fig. 4.6: Screen shot of GUI for SURPRISE

V. CONCLUDING REMARKS

The proposed system combines features of both ECG and thermal images in recognizing human emotions. Overall recognition rate using ECG is 61.9% and using ECG and thermal Images is 85.69%. Shortcoming of the system is lack of automatic identification of region of interest and smaller database. Contributing parts of the face which exhibited noticeable change in thermal signatures are cheeks, chins, and forehead. The performance of the work proposed can be improved by training the system with multiple images of same subject taken during different time intervals which may reveal few more contributing features and also by considering thermal images with larger resolution. Source of error may be due to thermal variability because of muscle deformation, in that case analysis of visual and thermal imagery can be combined which can yield larger accuracy [6].

ACKNOWLEDGMENT

Authors thank Malnad College of Engineering, Hassan and TEQIP III for supporting this research work.

REFERENCES

1. Ekman, P, “Face muscles talk every language.”, Psychology Today ,Vol.9, pp.35-39, 1975.
2. N. Fragopanagos, J.G. Taylor, “Emotion recognition in human–computer interaction”, Elsevier, Neural Networks 18, pp. 389–405, 2005.
3. Song, Kyu-Seob, et al. "Decision-Level Fusion Method for Emotion Recognition using Multimodal Emotion Recognition Information." 2018 15th International Conference on Ubiquitous Robots (UR). IEEE, 2018.
4. Hasan, Nahian Ibn, and Arnab Bhattacharjee. "Deep Learning Approach to Cardiovascular Disease Classification Employing Modified ECG Signal from Empirical Mode Decomposition." Biomedical Signal Processing and Control 52 (2019): 128-140.
5. Li, Cailong, et al. "Emotion recognition of human physiological signals based on recursive quantitative analysis." 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI). IEEE, 2018.
6. Goulart, Christiane, et al. "Emotion analysis in children through facial emissivity of infrared thermal imaging." PloS one 14.3 (2019): e0212928.
7. Gosavi, Deepali, et al. "Brain Tumor Classification Using GLCM Features and Neural Network." Available at SSRN 3372921 (2019).
8. Rusli, Nazreen, et al. "Mean of Correlation Method for Optimization of Affective States Detection in Children." IEEE Access 6 (2018): 68487-68497.
9. Kruger, G. H., Latchamsetty, R., Langhals, N. B., Yokokawa, M., Chugh, A., Morady, F., ... & Berenfeld, O. (2019). Bimodal classification algorithm for atrial fibrillation detection from m-health ECG recordings. Computers in biology and medicine, 104, 310-318.

AUTHORS PROFILE

Mr. C. M. Naveen Kumar obtained BE and M.Tech degrees from MCE, Hassan in 2010 and 2015 respectively. He is currently registered for part-time Ph.D programme under VTU, Belagavi. He is serving MCE since 2010 as Assistant Professor. He has attended 25 workshops/ Symposiums/conferences. His areas of interest are Digital Image Processing, ANN. He has published/presented 6 international journal papers, 4 international conference papers and 2 national conference papers.

Dr. G. Shivakumar obtained BE Degree in1990 from MCE, Hassan, M.Tech in 1998 from IIT Kharagpur, PGD(HRM) from KSOU Mysore in 2005 and Ph.D from VTU, Belagavi in 2015. He is serving MCE since 1990 and presently working as Professor & Dean(Research). He has attended 32 workshops/ Symposiums. . His areas of interest are Microprocessor and Microcontroller based System Design, Soft Computing, Pattern Recognition, Affective Computing and Virtual Instrumentation. He has published/presented 15 international journal papers, 20 international conference papers and 12 national conference papers.