ISSN: 0975-766X
CODEN: IJPTFI
Available Online through
www.ijptonline.com
Research Article

DETECTION AND ANALYSIS OF BIMODAL HUMAN EMOTIONS USING FPGA

Swagata Sarkar[1], H Ranganathan[2]
[1]Research Scholar, Department of ECE, Sathyabama University, Chennai-600119, India.
[2]Professor, Gojan School of Business and Technology, Chennai-600052 India.
Email: swagata.b.sarkar@gmail.com

Received on: 14-02-2017 *Accepted on: 22-03-2017*

Abstract:

Human emotion detection is very much relevant in today's scenario. There are six basic emotions. Those are happy, fear, anger, sad, surprise and neutral. It is very difficult to identify emotions separately. In this paper it is attempted to identify human emotion based on physical parameter changes and speech signal. The two different domain signals are fused by logistic regression. The classification is done by artificial neural network using back propagation algorithm. It is observed that sensitivity of bimodal emotion detection is 1% increased than single mode classification.

Keywords: Human Emotions, Physical parameters, FPGA, MFCC, Logistic Regression

Introduction

Human emotion detection is a very challenging task in today's scenario. Due to technological up gradation and high speed of life, people are under huge mental pressure which in turn generating lots of mental disorder. In this context it is very much necessary to develop an efficient human emotion detection system. In this paper bimodal human emotion detection approach is attempted. 14 physical parameters are collected from 50subjects of age range 21 to 23. The values are fuzzified o enhance the accuracy and combined with emotional speech parameters. Speech features are detected using Mel Frequency Cepstral Coefficients (MFCC) algorithm. 13 parameters are detected from MFCC algorithm. Three features are detected from FPGA analysis. All the 16 speech features are combined with 14 physical parameters using regression model. Finally, the features are classified using Artificial Neural Network (ANN) Back Propagation Algorithm (BPN).

Literature Review: In this paper both physical parameter features and speech signal features are fused by logistic regression. In reference [1] four physical parameters are taken as feature set and classified using K-Nearest

Neighborhood algorithm. 54.5% achieved for arousal and 38% achieved for valence. In [2], three physical parameters are considered. Random forest classifier is used to achieve 74% output. In paper [3] only one parameter is analyzed using artificial neural network and support vector machine (SVM) classifier to get 70% result. In reference [4] acoustic features are taken and classified by SVM classifier to get 73.56% output from SAVEE database. In reference [5] gesture and speech is combined to give the results as 76.28% and 85.39% respectively. Decision level fusion is used for the fusion of bimodal data. In reference [6] face and speech feature is combined to give 80% output using Boltzmann Zippers classifier with own database.

The flow diagram of the work follows in figure 1.

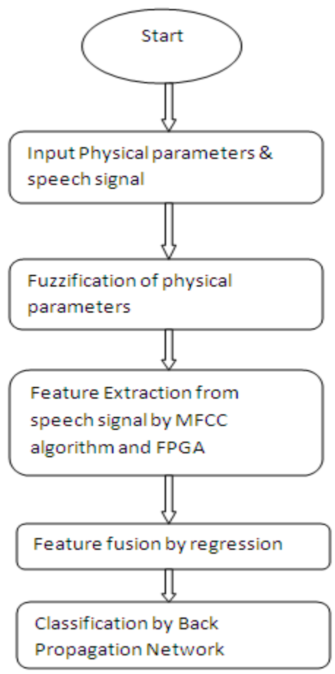

Fig.1. Flow Diagram.

The 14 physical parameters are directly related to the emotional changes of human being. The 14 parameters are tabulate as follows in table 1.

Table 1 Physical parameters affected by Human emotion change.

Sl. No.	Physical Parameter	Lower Range	Middle Range	Higher Range	Devices to measure
1	Electroencephalography (EEG)	alpha (13-15 Hz), beta (7.5-13 Hz), Theta (2.5-8), delta (<4 Hz			10-20 Electrode System

2	Heart Rate (HR)	20 to 70 BPM	45 to 100 BPM	84 to 120 BPM	Electrocardiogram Machine
3	Heart Rate Variability (HRV)	0.15 to 0.4 Hz	0.04 to 0.15 Hz	0.003 to 0.4 Hz	Electrocardiogram Machine
4	Pre-Ejection Period (PEP)	0 to 800 ms	0 to 1000 ms	500 to 1100 ms	Electrocardiogram Machine
5	Stroke volume (SV):	10 to 144 ml	10 to 250 ml	240 to 400 ml	Electrocardiogram Machine
6	Systolic blood pressure (SBP)	100 to 121 Hg	110 to 134 Hg	120 to 147 Hg	Blood pressure measurement system
7	Diastolic blood pressure (DBP)	77 to 87 Hg	81 to 91 Hg	85 to 91 Hg	Blood pressure measurement system
8	Skin Conductance Response (SCR)	0 to 0.2 ms	0.1 to 1 ms	0.85 ms to 1.5 ms	Skin Conductance Sensor
9	Tidal Volume (TV)	100 ml/breath to 150 ml/breath	200 ml/breath to 750 ml/breath	600 ml/breath to 1200 ml/breath	Spiro meter
10	Oscillatory Resistance (OR)	0 to 0.49	0.4 to 0.88	0.5 to 1	Spiro meter
11	Respiration Rate (RR)	5 to 10 breaths/minute	7 to 23 breaths/minute	15 to 24 breaths/minute.	Spiro meter
12	Non Specific Skin Conductance Response (NSCR)	0 to 2 per minute	1 to 3 per minute	2 to 5 per minute	Skin Conductance Sensor
13	Skin Conductance level (SCL)	0 to 2 ms	2 to 25 ms	20 to 25 ms	Skin Conductance Sensor

14	Finger (FT)	Temperature	65 ^{o}F to 75 ^{o}F	75 to 85 ^{o}F	80 to 90 ^{o}F	Finger Pulse Oximeter

Physical parameters are fuzzified for getting more specific results.14 input triangular membership functions are used with above ranges. 6 output membership functions are used for output emotions like happy, sad, fear, anger, disgust and neutral. Sugeno method is used for fuzzification and output of fuzzification is defuzzified by centroid method.

Speech signal is first preprocessed and then features are extracted by modified MFCC algorithm. The block diagram of MFCC algorithm is presented in figure 2.

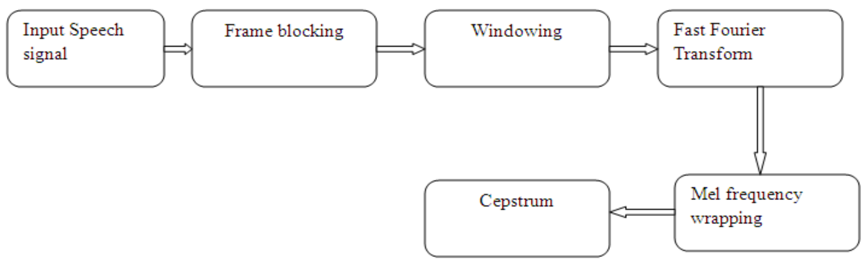

Fig.2.Block Diagram of MFCC algorithm.

Input speech signal is preprocessed and applied to the frame blocking. The output of frame blocking is fed to the windowing where different windowing techniques are used for efficient feature extraction. Emotion versus widowing table is given in table 2.

Table 2 emotion versus windows.

1. Sl. No.	**2. Emotion**	**3. Windows**
4. 1	5. Happy	6. Kaiser
7. 2	8. Sad	9. Bartlett
10. 3	11. Fear	12. Hamming
13. 4	14. Anger	15. Hanning
16. 5	17. Neutral	18. Hamming
19. 6	20. Surprise	21. Hanning

After windowing with different techniques outputs are fed to fast Fourier transform block. Output coefficients are taken from Cepstrum block. 13 coefficients are collected from MFCC algorithm.

Other three features are collected from FPGA implementations. LUTs, maximum and minimum asynchronous delays are the three parameters collected from FPGA output.

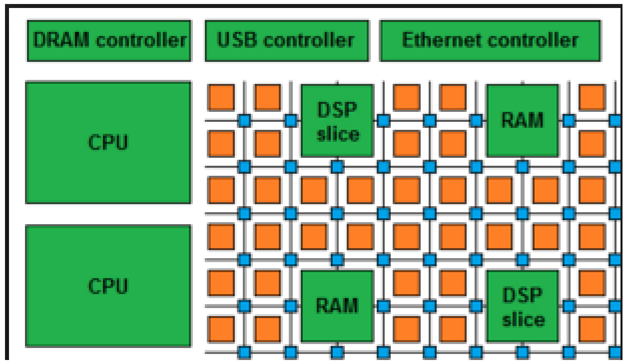

Fig.3. FPGA block.

The feature collected from physical parameters and speech signals are fused using logistic regression algorithm. In logistic regression the variables are categorical and it can model the non linear parameters.

Fused 30 features are classified using back propagation algorithm. There are three layers used in the back propagation algorithm. Those are input layer, hidden layer and output layer. Based on the number of features there are 30 input neuron in the input layer, 6 neurons in the output layer based on 6 different emotions output and 6 neurons are there in the hidden layer.

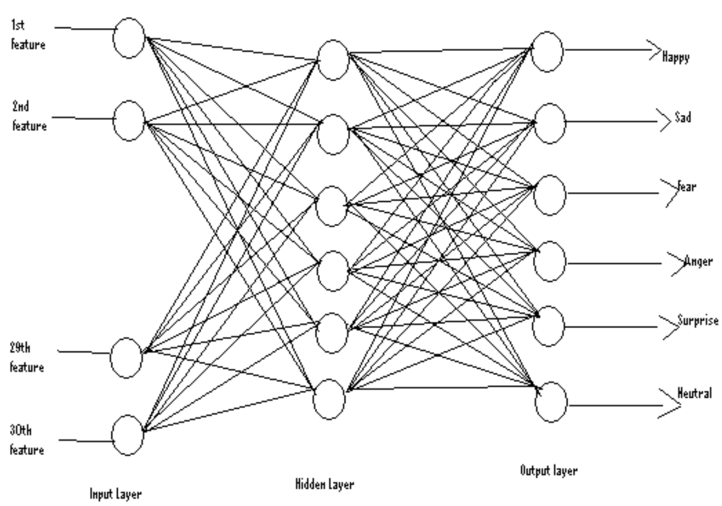

Fig.4. Architecture of ANN with 30 inputs & 6 outputs.

Results and Discussions

The results of different emotions can be analyzed by the specificity and sensitivity calculations.

Sensitivity

The sensitivity of a clinical test refers to the ability of the test to correctly identify those patients with the disease.

$$\text{Sensitivity} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

Specificity

The specificity of a clinical test refers to the ability of the test to correctly identify those patients without the disease.

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}}$$

The table 3 gives the output True Positive, True Negative, False Positive and False Negative for fuzzified physical parameters alone.

Table 3. Analysis of fuzzified physical parameters.

Sl. No.	Types of emotions	True Positive	True Negative	False Positive	False Negative	Specificity	Sensitivity
1	Happy	25	17	6	2	73.19	92.59
2	Sad	29	15	1	5	93.75	85.29
3	Fear	30	18	1	1	94.73	96.77
4	Angry	26	19	5	1	79.16	96.29
5	Neutral	25	20	3	2	86.95	93.10
6	Surprise	28	12	7	3	63.15	90.32

The same parameters can be calculated using fused features. The fusion is done for physical parameters with speech signals. Table 4 gives the analysis report for fused features.

Table 4. Analysis of fused features.

Sl. No.	Types of emotions	True Positive	True Negative	False Positive	False Negative	Specificity	Sensitivity
1	Happy	29	15	5	1	75	96.67
2	Sad	27	16	1	6	94.12	81.81
3	Fear	33	15	1	1	93.75	97.05
4	Angry	28	17	4	1	80.95	96.55

5	Neutral	27	18	3	2	85.71	93.10
6	Surprise	29	11	7	3	61.11	90.63

Figure 5 &6 show the comparison between specificity and sensitivity of physical parameters and fused features respectively.

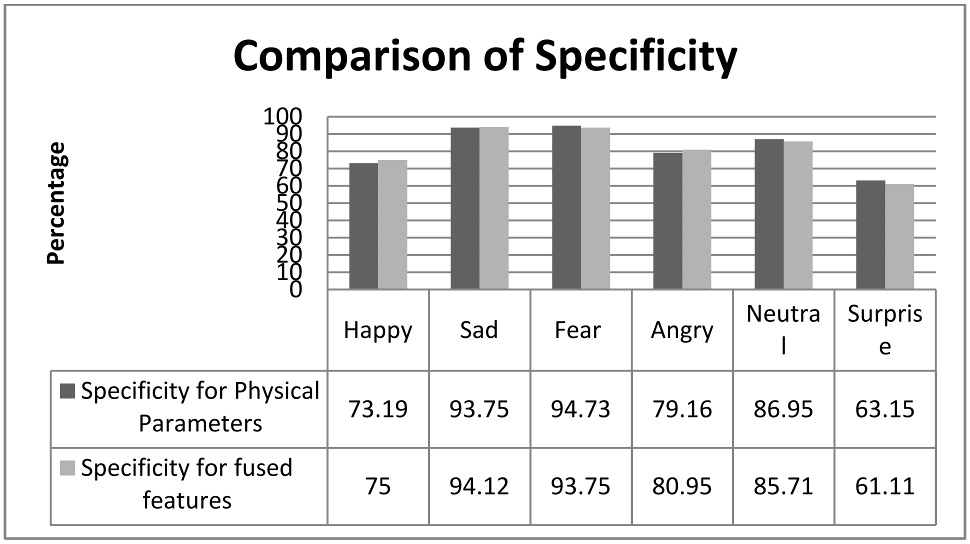

Fig.5. Comparison of Specificity.

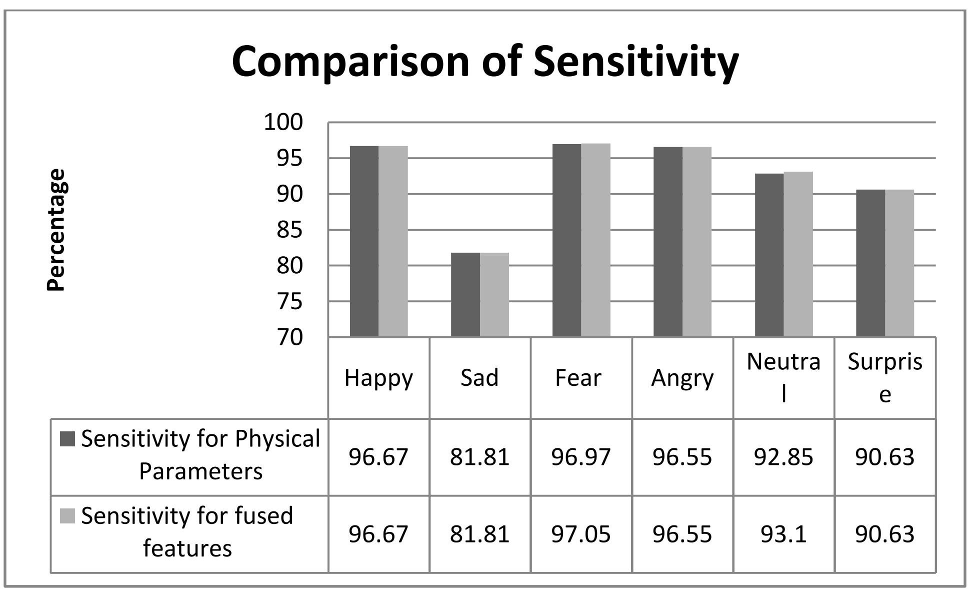

Fig.6. Comparison of Sensitivity.

In literature review it is seen that only few physical parameters are used for classification. Maximum 70% output is achieved. In this paper 14 physical parameters are used for classification. 81.82% specificity and 92.39% sensitivity are achieved. Fusing both speech feature and physical parameter, specificity and sensitivity become 82.77% and 92.65% respectively. The results are improved by 12% as compared with the literature.

Conclusions

In this paper comparison of sensitivity and specificity is done based on physical parameters and fused parameters. It is seen that fused feature classification gives 1% better sensitivity than single mode classifier.

Acknowledgement

The authors are like to thank the university to give wonderful research opportunity.

References

1. J. Kortelainen et. al. “Multimodal emotion recognition by combining physiological signals and facial expressions: a preliminary study”, in proc. Of IEEE Int’l conference on embedded Systems, pp. 5238-5241, Sep. 2012. DOI 10.1109/EMBC.2012.6347175.
2. D. Lakens, “Using a smart phone to measure heart rate changes during relieved happiness and anger”, IEEE Transactions on Affective Computing, vol. 4, pp. 238-241, 2013. DOI 10.1109/T-AFFC.2013.3.
3. Wanhui Wen et. al., “Emotion recognition based on multivarient correlation of physiological signals”, IEEE transaction on Affective Computing, vol. 5, pp. 126-140, 2014. DOI: 10.1109/TAFFC.2014.2327617.
4. Carlos B, Mariooryad S, Metallinou A and Narayanan S “Iterative feature normalization scheme for automatic detection from speech”, IEEE transaction on Affective Computing, vol.4,pp.386-397,2013,DOI.10.1109/T-AFFC2013.26.
5. H.A. Vu, y. Yamazaki, F.Fong and K. Hirota, “emotion recognition based on human gesture and speech information using RT middleware”, In proc. of IEEE int’l conference on fuzzy system, pp. 787-791, 2011.DOI. 10.1109/FUZZy.2011.6007557.
6. K.Lu and Y. Jia, “Audio visual emotion recognition using Boltzmann zippers”, in proc. of IEEE Int’l conference on image processing, pp. 2589-2592, 2012, DOI 10.1109/ICIP.2012.6467428.