Determination of Angry Condition based on EEG, Speech and Heartbeat

Masnani Mohamed[1], Lee Ri Quan[2], Ida Laila Ahmad[3], Lee Chee Chuan[4] and Siti Hanira bt Hamid[5]
Faculty of Electrical and Electronic Engineering
University Tun Hussein Onn Malaysia
86400 Parit Raja, Batu Pahat, Johor, Malaysia
[1]masnani@uthm.edu.my, [2]riquan89@gmail.com, [3]ida@uthm.edu.my, [4]charles_lee1231@hotmail.com, [5]me.nira89@gmail.com

***Abstract*—This paper determines the angry emotion condition by analyzing and recognizing speech signal, EEG signal, as well as detecting the heartbeat. For the speech analyzing experiment, several digital signal processing methods such as autocorrelation and linear predication technique was introduced to analyze the features. Then, Artificial Neural Network (ANN) was used to classify each parameter features such as mean fundamental frequency, maximum fundamental frequency, standard deviation fundamental frequency, mean amplitude, pause length ratio and first formant frequency to recognize the emotion. For the EEG analysis, the raw EEG signal was undergone preprocessing to remove the artifacts to minimal. Some features as mean, standard deviation, the peak amplitude, the peak amplitude in alpha band (PAA) and the peak frequency in alpha band (PAF) of the EEG signals were extracted. The selected features were classified by using ANN to obtain the maximum classification accuracy rate. Meanwhile, a heartbeat monitoring circuit was developed to measure the heartbeat. The result showed that angry emotion has relatively low condition in mean value, maximum peak amplitude and relatively high peak frequency in alpha band (PAF) of the EEG signals. The mean fundamental frequency, standard deviation fundamental frequency and mean intensity of the speech signal are good in determining the angry emotion. This method can be used further to recognize angry emotion of patient during counseling session.**

Keywords-Artificial Neural Network, digital signal processing, Electroencephalogram (EEG), emotion recognition, emotional speech signal and heartbeat.

I. INTRODUCTION

Emotion is a complex psycho-physiological experience of an individual's state of mind as interacting with biochemical (internal) and environmental (external) influences. There are some models proposed in order to classify and represent emotions. There are models use the idea that all emotions can be composed of some basic emotions, just as colors can be composed of primary colors [1, 2]. For both theoretical and practical reasons some researchers define emotions according to one or more dimensions. One of the popular versions is dimensional or circumflex model, which uses the models of arousal and valence as shown in Fig. 1[3]. The third dimension seldom used in research, it is not always the same even used. Dominance and motor activation (approach or avoid) are ever used as the third dimension [3].

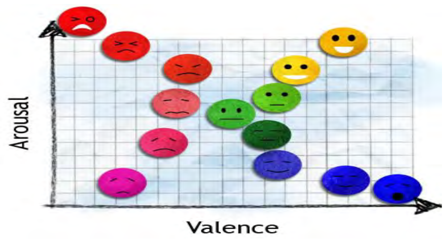

Figure 1. Different emotions on valence and arousal scale.

There are few methods can be used to identify the type of emotions. For example by using the EEG and speech signal analysis, the particular state of emotion can be determined. In addition, the heartbeat variations also can help to monitor the state of the individual's emotion. The EEG signal typically described in term of rhythmic activity and divided into bands by frequency. The EEG signal primarily categorized into 5 different frequency bands with each normally rhythmic activity. Table 1 shows the comparison table for each different frequency band. The pattern of each frequency bands signal is shown in Fig. 2.

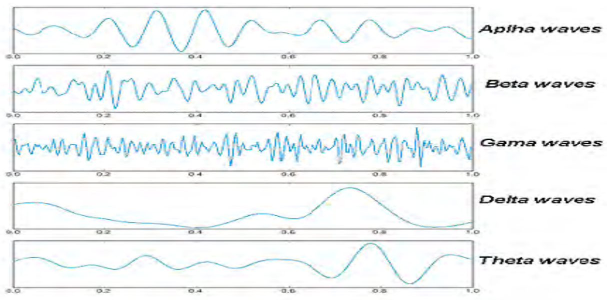

Figure 2. EEG pattern.

TABLE 1. RHYTHMIC ACTIVITY FOR EACH EEG FREQUENCY BANDS

Type	Frequency, f (Hz)	Activity
Delta	0 ≤ f < 4	Deep sleep or unconscious
Theta	4 ≤ f < 8	Drowsiness, meditation, recall, dream or intuitive
Alpha	8 ≤ f < 13	A person is awake, and is known to be apparent when eyes are opened, relaxed or conscious
Beta	13 ≤ f < 30	Active thinking or concentration
Gamma	30 ≤ f < 100	Displays during cross-modal sensory processing

The aim for the EEG signal analysis is to focus on comparing the difference between human unpleasant emotion (disgusting and angry) with human pleasant (happy and surprise) and neutral emotion. An open EEG signal database, Enterface'06 EEG signals database was used as part of this study. The raw EEG signals were preprocessed to remove the artifacts to minimal. Different methods of feature extraction techniques were presented. The mean, standard deviation, the peak amplitude, the peak amplitude in alpha band (PAA) and the peak frequency in alpha band (PAF) of the EEG signals were extracted. The selected features were classified by using Artificial Neural Network to obtain the maximum classification accuracy rate.

Many researchers prove that assessing emotions is a key to understand human nature, especially for angry emotion. Anger is a state where a person feels angry, mad and exasperates to a certain condition. One of the methods proposed in this paper is to analyze and recognize the speech signal because it is the most important media of human interaction contains a lot of emotional information [4]. The pitch, tone, timing and energy of speech are all jointly influenced in a nontrivial manner to express the emotional message [5]. For this part, emotion is roughly classified into two categories which are angry and no angry emotion. The no angry emotion consists of happy, normal and sad speech signal. However the main focus is on angry speech signal. Every emotional speech has its own features, such as fundamental frequency, amplitude, formant frequency, duration and so on. Therefore this project is used to analyze these features and compared the two emotions speech signal in order to obtain the pattern of distribution of different speech.

For the analyzing experiment, in terms of sentences length, environment issue and gender differences are all considered. Generally several methods such as autocorrelation and linear predication technique have been introduced in order to analyze the features, subsequently classified such as mean fundamental frequency, maximum fundamental frequency, minimum fundamental frequency, standard deviation fundamental frequency, mean amplitude, mean first formant and pause length ratio using Artificial Neural Network in order to obtain relationship between input parameters and output emotions. Eventually a MATLAB GUI has been created by displaying the feature parameters of angry speech signal as well as recognizing the emotion.

It has been widely studied and analyzed that heart rate can be measured by the number of heartbeats per unit of time; an accessible parameter that can be acquired easily [6]. As the heart is beating, an electrical signal is transmitted through the heart muscle and during this moment, the heartbeat can be detected by using infrared sensor. The sensor is placed at fingertips and infrared light is emitted into skin. Basically, the light absorbed is depending on the blood volume [7].

II. METHODOLOGY

A. EEG Signal Analysis

Fig. 3 shows the flowchart of the project to give brief understanding on this work.

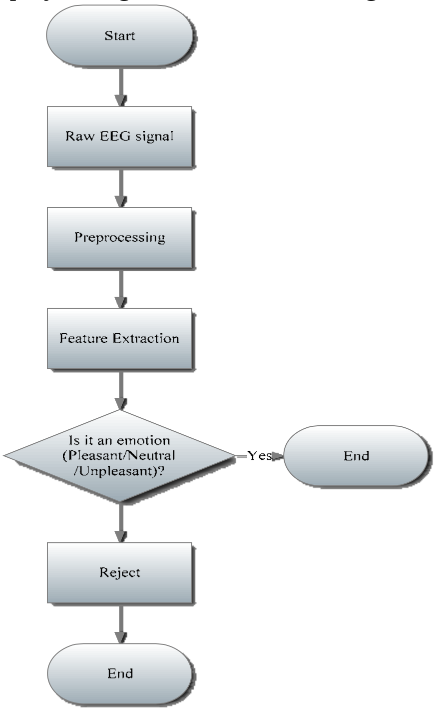

Figure 3. Structural overview of emotional recognition system protocol.

1. EEG Database

For this study, the EEG signals database from Enterface'06 (referred to as Enterface database) was used in this project [9]. Enterface experiment was meant to provide a database of EEG and other physiological signals for emotion recognition to public use, copy and publish.

Each participant was gone through three sessions of 30 emotional stimuli in Enterface experiment. Visual stimuli were used in experiment to elicit the participant emotion, in particular images from International Affective Picture System (IAPS). Each emotional stimulus consists of five same class images to ensure stability of the emotion over time. Every image was displayed for 2.5 seconds leading to a total 12.5 seconds for each emotional stimuli block. A 10 seconds dark screen to participants calm down their emotion. In the mean while, the participant was asked to self-assess their emotional state. These self-assessments are distributed because emotions are known to be very subjective dependent. One can never be sure that the person emotion after viewing the images and this self-assessment can be used as evaluation for different participant. Fig. 4 shows the Enterface experiment protocol.

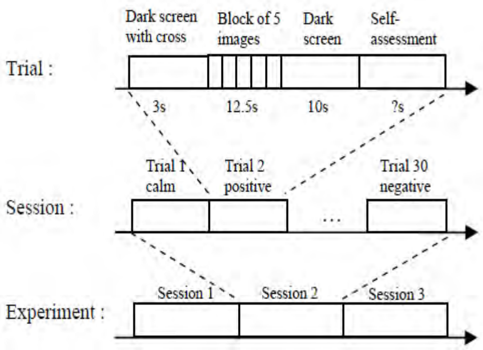

Figure 4. Enterface experiment protocol description.

Basically, each participant was having thirty trials in each session, and each session have involved with five participants. One single session was consisted of 150 samples. The overall data was consisted of 450 samples.

2. *Channel Selection*

The EEG signals database was recorded data from participants using the Biosemi Active 2 acquisition system with 64 EEG channels and the peripheral sensors. Until today, the researchers still cannot find the specific region on skull where the brain activity is sufficiently high to detect an emotional state. Therefore, 15 channels were collected arbitrary and 1 channel was selected as the reference channel. The channels AF3, AF4, F3, F4, FCz, C3, C4, CPz, P3, Pz, P4, POz, O1, Oz and O2 were selected for the feature extraction. The reference channel is Cz.

The increased of the number of channels from 2 channels to 16 channels will increase the result of maximum accuracy classification. However, it can also be noted that testing classification starts to decline after selecting more than 20 channels, and decline become more rapid when selecting 40 or more channels. The declination occurred due to more parameters for classifier to estimate, which make it harder to generalize, especially when dealing with limited number of patterns [8].

Fig. 5 shows the maps of the placement of the electrodes according to the “10-20 electrodes placement system”. The green colour circle channels are the selected channel and the red colour circle channel is the reference channel.

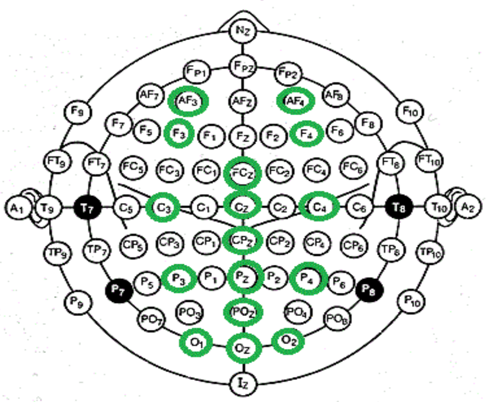

Figure 5. Channel selected.

3. *Preprocessing*

Preprocessing is the stage of processing raw EEG signals such as removing noise, artifacts, and other interferences. The measured raw EEG signal is a combination of activity, reference activity and noise as shown in Fig. 6. All the process was done by using EEGlab which is an additional toolbox in Matlab software as shown in Fig. 7.

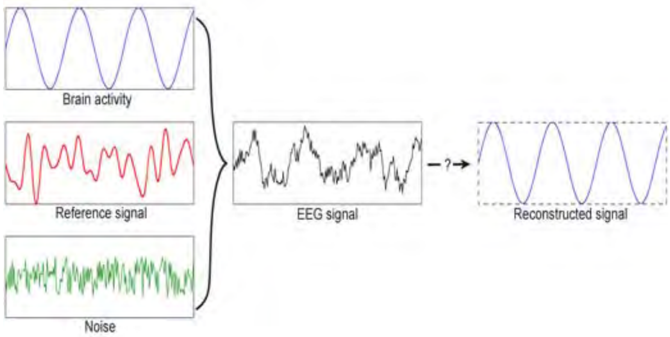

Figure 6. An EEG signal is a combination of brain activity, reference activity, and noise.

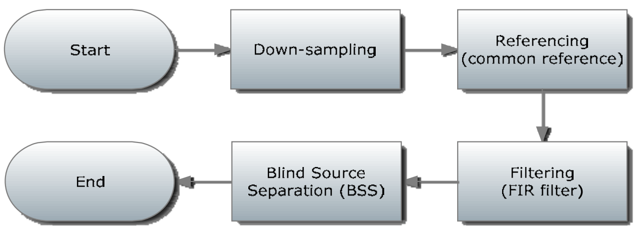

Figure 7. Block diagram of preprocessing protocol.

a. Down-sampling

Down-sampling is the process of reducing the sampling rate of the EEG signal to reduce the date rate or the size of the data. This process can save Matlab memory space and preserving all the information. The original raw EEG data is measured at the rate of 1024 Hz was reduced to 512 Hz.

b. Referencing

The differences between results of different studies are partly due to the differences in referencing [9]. The acquisition of EEG signals is a recording of electrical activity originating at different electrodes on the human being skull. Since the brain activity voltage is a relative measure, the measurement is compared to the voltage at reference site. This results in a combination of brain activity at the given electrode, brain activity at the reference site and noise. Because of this, the reference site should be chosen such that the brain activity at that site is almost zero.

The common reference technique will be applied as method of referencing. This method is applied in the experiment widely used by many researchers. The configuration with 8 electrodes gives the best result for the studying data as shown in Fig. 8 [10]. The linked mastoid (the vertex electrode (Cz)) is used as a reference electrode.

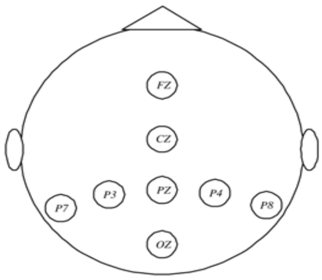

Figure 8. The references channels with 8 electrodes given by Hoffman et al. (2008).

c. Filtering

The raw EEG data normally contain a lot of noises and artifacts. Some examples of noise sources are within 50 and 60 Hz power line interference, fluorescent lighting, baseline drift (low frequency noise), electrocardiogram (ECG), electromyogram(EMG), and random noises. Therefore, a Finite Impulse Response (FIR) filter will be designed to filter the frequency lower 2 Hz and greater than 45 Hz. There is not much brain activity with very low frequencies, most of the signal is artifacts [11]. The signal above than 40Hz is a high frequency noise. The FIR filter is applied in the project instead of using Infinite Impulse Response (IIR) filter because they are more stable, and feedback is not involved. Fig. 9 shows the signal pattern after filtered.

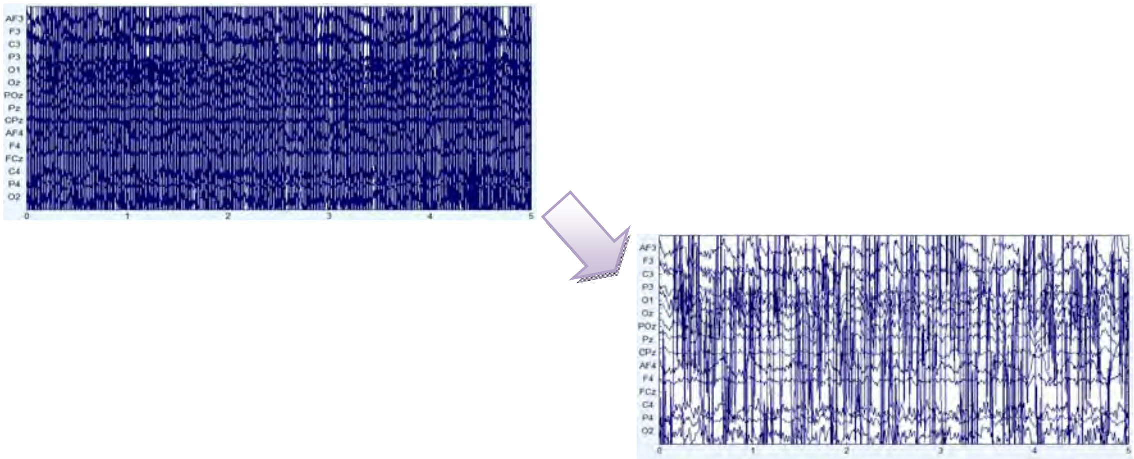

Figure 9. The filtered EEG signal.

Blind Source Separation (BSS) technique is used to remove electroocular (EOG) and muscular (EMG) artifacts. Fig. 10 shows the EEG signals pattern after applying BBS technique. Actually, the signal does not has much changes after applying BSS technique.

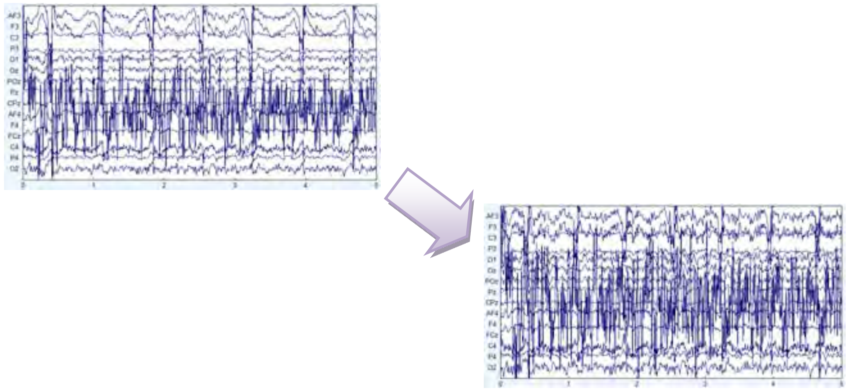

Figure 10. EEG pattern after BBS technique applied on EEG data.

BBS does not need separate EOG measurement since it is a separation process on a set of signals from a set of mixed signal, without the aid of information (or with very little information) about source signals or mixing process. BSS relies on the assumption that these independent sources indeed exist. For EEG signals, this assumption holds. As the EEG signals on an electrode are said to be a mix of underlying signals of brain processes, the signals can be decomposed. An example of BSS is shown in Fig. 11, where all different samples can be coarsely separated into two groups, corresponding to the blue and the red line [11].

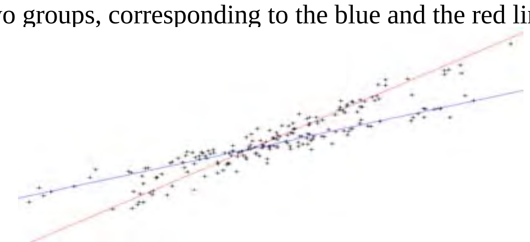

Figure 11. Example of BSS. Two components are found within the data.

Anyway, that is impossible to remove all the artifacts in EEG signals. The complete removal of artifacts will also remove some useful information of EEG signals [12].

4. Feature Extraction

Feature extraction is the process of extracting useful information from the EEG signals. Features are characteristics of signal that are able to distinguish between emotions. Feature selection is an important step for the final result of the classification. Matlab software is used as an intermediate to extract the EEG features. All the features were extracted from fifteen channels for each signal or trial.

a. Mean Value

Each sensor channel represented x contains particular number of value. Mean value of each channel is used as a feature.

$$\mu_{channel} = \frac{1}{N}\sum_{i=0}^{N} X_i \qquad (1)$$

where N represents window length or sum of the sensor channel.

b. Variance

Variance for each channel also used as a feature selection.

$$\sigma_{channel} = \frac{1}{N}\sum_{i=1}^{N} (X_i - \mu_{channel})^2 \qquad (2)$$

where N represents window length.

c. Maximum and Minimum Amplitude

Maximum and minimum amplitude for each channel is determined and recorded.

d. Peak Amplitude and frequency in Alpha Band (PAA & PAF)

Alpha frequency band lies between the range 8-13 Hz. These waves have comparatively higher amplitude to other wave types. Because this reason, alpha wave type is useful to recognice mental and/or emotion effort. In order to obtain the alpha frequency band signal, a Finite Impulse Response (FIR) bandpass filter is needed. After the Fast Fourier Transform (FFT) was applied on the PAA signal to transform the signal from time domain to frequency domain. Then, found out the maximum value between this frequency range.

5. Classification

The Artificial Neural network (ANN) technique will be used to classify each EEG features into three basic emotions which is pleasant, neutral and unpleasant. Neural network detection systems have been proposed by a number of researchers [7, 11]. The neural network pattern recognition tool in the Matlab program is used to solve a pattern-recognition problem with a two-layer feed forward network.

Artificial neural networks are computing systems made up of large number of simple, highly interconnected processing elements (called nodes or artificial neurons) that abstractly emulate the structure and operation of the biological nervous system. The ANN classifier required one set of inputs data and one set of targets data with the condition both data must have a same sample size. In this project, an input data is the feature extracted from EEG signal and three states of emotion treat as a target data. Sixty percent of input data is used as training data, another twenty percent of input data is used as testing data and the rest of the input data is validation data. The target output for each emotion is set as '100' for pleasant, '010' for neutral and '001' for unpleasant.

Validation data are used to measure network generalization, and to halt training when generalization stops improving. Testing data have no effect on training and so provide an independent measure of network performance during and after training. Training data is the data used to train in the network.

B. Speech Signal Analysis

Fig. 12 shows the block diagram of the feature analysis and recognition.

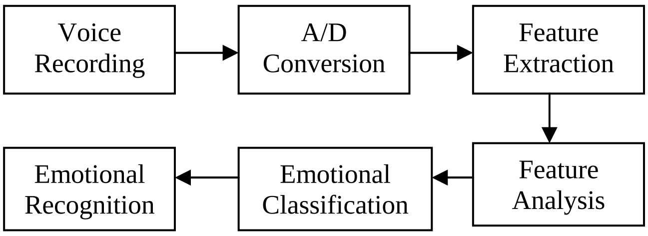

Figure 12. Block diagram of feature analysis and recognition.

1. Voice Recording

The recording of the voices through microphone were in a quiet room in order to reduce the environment noise. All speakers in different gender uttered the same sentences in different content (angry and non-angry) in order to allow the comparability across emotions and speakers. In order to test the validity of emotion data collected from this experiment, a listening pre-experiment is launched; all speakers are required to tell the emotion type of the sentences and tested by listener.

2. Analog to Digital Conversion

The analog-to-digital converter (ADC) translates this analog wave into digital data that the computer can understand. Voice was undertook A/D conversion into digital signal through personal computer by voice recorder software at 44 kHz sampling rate and saved into Wave file. After that, the entire speech signal in WAV file was analyzed by MATLAB.

3. Feature Extraction and Analysis

Features of the sampled signal such as: pitch, formant frequency, amplitude and duration were extracted individually by using different algorithm. The first stage of analysis process is windowed the speech signal by a hamming window with length of 30 ms and window was shifted to 15 ms so that the overlapping will occur in between windowed [14]. The four selection features are shown in next stages which are pitch, formant frequency, amplitude and duration.

a. Pitch Detection by Autocorrelation Approach Form

Autocorrelation requires that several initial assumptions be made about the set or sequence of speech samples. First, it requires that signal must be stationary and second, it requires that the sequence is zero outside of the current segment. Therefore, by using the windowed signal (stationary and zero outside of current segment), short time autocorrelation has been adopted in order to detect fundamental frequency tone frame by frame.

b. Formant Detection by Linear Predication Approach

Normally, vocal tract resonances will cause peaks in the observed spectral envelope. Therefore, linear predication analysis was used to find the best matching system by passed through a purely-recursive IIR filter. In order to find formant frequency from the filter, it is necessary to find out the locations of the resonances that make up the filter. This involves treating the filter coefficients as a polynomial and solves the roots of the polynomial. Normally first and second average formant frequency of every frame was used as features parameters [5].

c. Amplitude

The short-term speech amplitude has been exploited for emotion recognition, because it was related to the arousal level of emotions. The raw data that plotted by time domain was used to analyze amplitude of speech signal. The amplitude shown in time domain graph is energy of the speech signal since the parameter needed to recognize emotion is intensity. Therefore, the energy was converted into intensity. After the intensity of each windowed signal has been calculated, the parameter such as mean intensity of each sentence was taken as features parameter. The reason of calculating the intensity of each windowed signal is to identify the unvoiced signal which may decrease the accuracy of intensity.

d. Duration

Total uttered duration and total pause duration between speech segments of every emotion sentence has been obtained. By analyzing short term windowed signal, the pause duration between speech segments was calculated. Several assumptions have been considered in order to calculate pause duration such as: low energy, high frequency and low autocorrelation peak [15]. In recognition, the ratio of total uttered duration divided by pause duration of selected sentences was taken as feature parameter.

4. Classification and Recognition

After extracting the desired features, the emotional classification technique was needed to train and test the data. After that, emotions have been recognized. In this project, Artificial Neural Network (ANN) was used to classify each parameter features including mean, maximum, minimum and standard deviation from fundamental frequency, mean amplitude, average sentences duration and first formant frequency and recognize the emotion.

C. Heartbeat Detection

The heartbeat is measured using the pulse detector circuit as shown in Fig. 13. It has two operational amplifiers (Op-Amps) functioned as non-inverting amplifiers that used to increase the input voltage. The input signal is connected to pin 3 of the LM358 and then amplified by both of amplifiers used. The high intensity LED used to produce a light towards the LDR. If the intensity of the light is strength, the resistance of LDR will be low while if the light intensity is weak, the resistance of LDR will be high. When a finger is placed between the high intensity LED and LDR, the light intensity will be reduced.

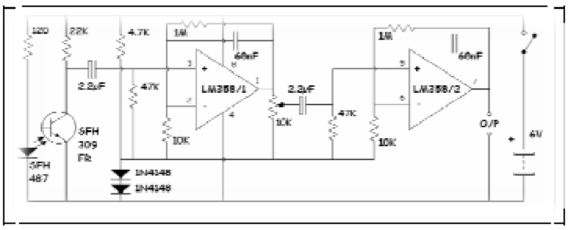

Figure 13. Pulse detector circuit.

III. RESULT AND ANALYSIS

A. EEG Signal

The maximum classification accuracy for the three sessions of the experiment had achieved as shown in Fig.14. They were trained by using ANN classifier. Table 2 shows the relationship between three emotions and the six extracted features.

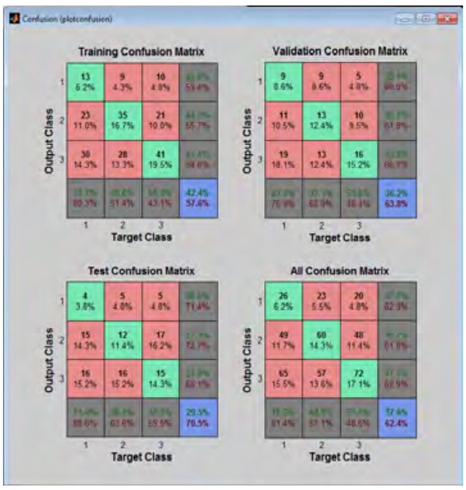

Figure 14. Overall maximum classification accuracy result.

TABLE 2. RELATIONSHIP BETWEEN EMOTIONS AND FEATURES

Emotion	Pleasant	Neutral	Unpleasant
Mean	High		
			Low
Variance		High	
		Low	
Minimum Peak Amplitude		High	
	Low		
Maximum Peak Amplitude		High	
			Low
PAA	High		
		Low	
PAF			High
		Low	

From the Table 2, it clearly shows that the EEG features trend of unpleasant emotion compare to another two emotion. The unpleasant emotion comes with a relatively low mean value of the EEG amplitude and maximum peak amplitude. However, an unpleasant emotion achieves peak amplitude at relatively higher frequency within the alpha band.

B. Speech Signal

For the result and analysis, all the feature parameters of speech signal have been distributed as database of the project. Subsequently, by using Neural Network Pattern Tool from MATLAB, the relationship between input parameter and output emotions were obtained by classified input into set of target categories.

1. Features Parameters Database Distribution

For the acoustic analysis of speech database the features parameters were chosen as representative part of database. By using MATLAB, all the feature extraction method such as autocorrelation and linear predication was applied in order to extract the related parameters features for male and female speakers. Table 3 shows the emotional speech information for 20 speech databases. The speakers were given a sentence to utter with two different emotions which are angry and no angry. Subsequently the recorded speech was analyzed and extracted through MATLAB by obtaining the important features parameters such as mean pitch, maximum pitch, minimum pitch, standard deviation pitch, mean intensity, mean first formant frequency and pause length ratio as in Table 3.

2. Classification using Neural Pattern Recognition Tool

All the features parameters obtained in Table 3 have been used as input databases while emotions were used as targets. A two layers feed-forward network with sigmoid hidden and output neuron was classified arbitrarily well with given enough neurons in its hidden layer. In this case, the number of hidden neuron was set at 20, which is considered as desired number of hidden neuron.

During the classification process, input database was divided into three kinds of samples, which are training sample, validation sample and testing sample. Training samples are presented to the network during training, and the network is adjusted according to its error. While validation samples were used to measure network generalization, and to halt training when generalization stops improving. At the meantime, testing sample provide an independent measure of network performance during and after training due to there have no effect on training. Subsequently, 70% of input data was used as training samples, 15% of input data is for validation samples and 15% input data for testing samples. This percentage considered as desired for classification. After that, network was trained until the satisfied result of all confusion matrixes has been obtained.

Performance of classification by using confusion matrix was divided into three main parts which are training confusion matrix, validation confusion and test confusion matrix, while all confusion matrixes was used to evaluate the overall performance. Inside the confusion matrix, output class and target class played the role in performance of classification. Target class is the emotion output that has been inserted according to input data. Meanwhile the output class is the result of emotional classification.

TABLE 3. EMOTIONAL SPEECH INFORMATION

G	Mean Fo (Hz)	Max Fo (Hz)	Min Fo (Hz)	SD Fo (Hz)	Mean Intensity (dB)	Mean F1 (Hz)	P/L	E
M	136	154	119	26.5	83.8	361	0.12	N
M	134	160	123	37.0	85.3	332	0.17	N
M	164	251	133	49.5	83.5	457	0.54	A
M	244	398	145	34.6	85.7	425	0.13	A
M	212	372	153	29.3	82.6	360	0.09	A
M	205	286	132	38.7	80.0	418	0.27	A
M	154	176	128	15.3	71.6	440	0.40	N
M	117	122	109	2.31	73.2	268	0.13	N
M	213	302	163	25.8	76.5	423	0.12	A
M	111	127	104	15.2	78.8	415	0.58	N
F	228	302	208	15.2	78.4	330	0.16	N
F	257	301	232	27.9	75.1	448	0.22	N
F	350	432	258	34.1	73.2	500	0.16	A
F	241	274	227	13.1	79.4	400	0.21	N
F	259	302	228	35.6	75.4	464	0.29	N
F	308	483	276	67.5	73.1	593	0.17	A
F	264	302	167	58.8	75.5	389	0.17	N
F	403	456	278	29.0	79.1	485	0.30	A
F	201	254	188	18.7	71.8	395	0.33	N
F	336	453	289	21.1	74.5	496	0.11	A

Explanation of Symbols: Max: Maximum, Min: Minimum, F0: Fundamental Frequency,
SD: Standard Deviation, F1: First Formant, P: Pause, L: Length, E: Emotions, A: Angry, N: Non-angry.

Fig. 15 shows the confusion matrix of male speakers and Fig. 16 shows the confusion matrix of female speakers. Inside both confusion matrix figures, the number 1 represented angry emotion and number 2 represented no angry emotion, (shown at horizontal and vertical axis of each matrix boxes).

Fig. 15 shows the performance confusion matrix of male speakers, which gives details of the strengths and weaknesses of this system. All confusion matrixes show the results for training, validation and testing experiment which are up to 86.5%. Means that the input parameter obtained have very good relationship with the emotions. Eventually this result was used as reference by recognize emotion in real time or loaded existing emotional speech signal.

Figure 15. Confusion matrix of male speakers.

Fig. 16 shows the performance of confusion matrix of female speaker. All confusion matrixes show 98.1% of performance classification. This can be concluded that the input parameter have very good relationship between output emotions.

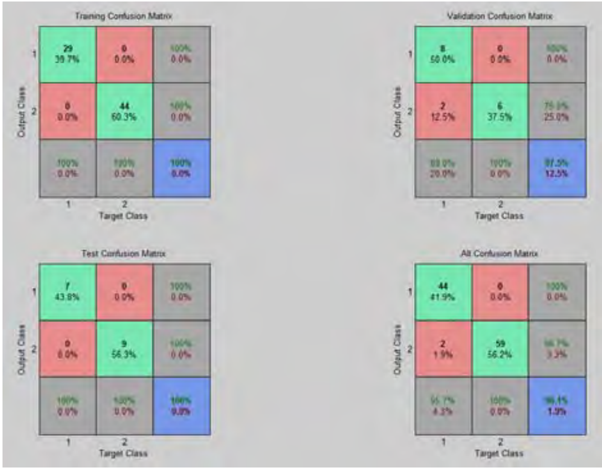

Figure 16. Confusion matrix of female speakers.

3. Emotional Recognition by using MATLAB GUI

MATLAB GUI has been build (Fig. 17) in order to recognize the emotional speech signal. By applied the feature extraction methods which are short time autocorrelation, linear predication coding and time domain presentation, all the features parameters of speech signal have been displayed in MATLAB GUI. After that, result of classification for Neural Pattern Recognition Tool was employed by recognizing the emotion from speech signal.

GUI has several functions, the graph axes were used to display the waveform of speech signal, for the data input section, there are 3 buttons to be chosen. First is "Female" button, which is used to load the existing female speech signal. Second is "Male" button, used to load the existing male speech signal and last is "Record" button, which is used to record the sound of speaker in real time. At the meantime, the feature parameter section shows the output features parameters of speech signal which are mean pitch, standard deviation pitch, maximum pitch, minimum pitch, mean intensity, mean first formant frequency and pause length ratio. The last section is emotion recognition, which is used to display the emotions of speech signal.

In order to test the emotional speech, first, press the desired button in order to choose speech signal. Once the desired speech file has been chosen and loaded, the waveform of speech signal was shown in the graph axes. Subsequently the features parameters of the speech signal were displayed at the feature parameter output. Finally the emotion was detected and displayed at emotion recognition section.

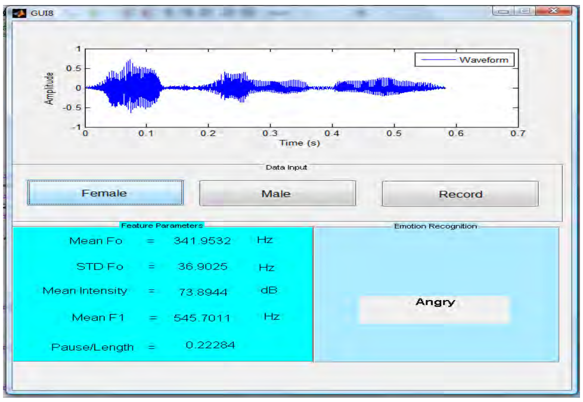

Figure 17. MATLAB GUI for speech recognition.

4. Result of Emotion Recognition

The results of emotion speech recognition have been tested by using existing file and in real time. The "real emotion" meaning the actual emotion of speech signal, whereas the "recognized emotion" is the emotion that has been tested by the system. The "result" shows the outcome of each speech signal that has been tested, if the actual emotion of speech signal match the recognized emotion, then the result will shows "correct", otherwise it will show "wrong".

The result of emotion recognition by loading existing file has been achieved 80% of accuracy. However, emotion recognition for real time is only 70% of accuracy.

C. Hearbeat Result

Table 4 shows the heartbeat taken whenever the speakers uttered the given sentence with two different emotions which are angry and no angry. It shows that when a person is angry, the heartbeat goes higher than normal. In this case, 90 bpm and above are considered as high.

TABLE 4. HEARTBEAT DETECTION

Subjects	Heartbeat (bpm)	Emotion (ANGRY/NORMAL)
1.	89	NORMAL
2.	86	NORMAL
3.	90	ANGRY
4.	90	ANGRY
5.	97	ANGRY
6.	90	ANGRY
7.	84	NORMAL
8.	72	NORMAL
9.	92	ANGRY
10.	88	NORMAL
11.	70	NORMAL
12.	73	NORMAL
13.	91	ANGRY
14.	88	NORMAL
15.	84	NORMAL
16.	98	ANGRY
17.	78	NORMAL
18.	102	ANGRY
19.	88	NORMAL
20.	96	ANGRY

IV. DISCUSSION AND CONCLUSION

A. Discussion

1. EEG Signal

Emotion Recognition from EEG signals still very young and a challenging task. There is a lot of problem need to overcome. One of the challenges of this project was to work with big database. For the experiment, many participants complained that feeling headache at the end of each experiment session. Some participants reported that the effects of the images decrease after viewing many images [13].

From the results obtained, the maximum classification accuracy rate that combined all the three sessions features ware only achieved 37.6%. This problem occurred due to the increasing of the number of parameters for the classifier to estimate will make it harder to generalize [8]. Each participant was having thirty trials in each session, and each session have involved with five participants. One single session was consisted of 150 samples. The overall data was consisted of 450 samples

Besides, there are still few reasons will lead to overall maximum classification accuracy rate for this research only obtained 37.6%. First, the features extracted from the signal are lack of effective on differentiate human emotion. Second, the channel selected in this research cannot show the different of emotion state clearly. Third, according to a researcher who was using the same EEG signals database from Enterface'06 stated that the Enterface's EEG signals database was not suitable for recognizing the human emotion [12].

The k-Nearest Neighborhood (KNN) classifier and LDA (Linear Discriminant Analysis) classifier were used to classify the emotion in two dimensional models, which were arousal and valence. The classification accuracy rate obtained for KNN and LDA classifier were 51% and 40% respectively. Therefore, the overall maximum classification accuracy result obtained in this research is still acceptable. The different of the result was because the different of the classifier.

2. Speech Signal

Another goal of this paper is to extract, analyze, classify and recognize the angry speech signal. Feature extraction and analysis play a most crucial role [16]. In this experiment, feature parameters has been extracted successfully by using different method of digital signal processing, such as, windowing, framing, autocorrelation and linear predication coding.

Classification also shows an outstanding result by employed the neural pattern recognition tool which has been achieved 86.5% for male classification and 98.1 % for female classification. The purpose of separated female and male classification is because, male and female has significant difference in their vocal size means that they have differently pitched voices. Moreover, men generally speaking, have a larger vocal tract, which essentially give the resultant voice higher compared to women. When a person in anger state he will speak faster with bigger volume and raised the tone due to produced changes in respiration and an increase in muscle tension, which influence the vibration of the vocal folds and vocal tract shape, affecting the acoustic characteristics of the speech.

B. Conclusion

The selected features of EEG signals were classified by using Artificial Neural Network (ANN) application to obtain the maximum classification accuracy rate. The research result showed that a person in unpleasant emotion have relatively low condition in mean value of the EEG signals and maximum peak amplitude EEG channels compare to others two emotions. However, a person is in unpleasant emotion when the peak frequency in alpha band (PAF) of the EEG signals is relatively high. In conclusion for EEG analysis, the mean value and the maximum peak amplitude value of the EEG signal are indication of an unpleasant EEG signal emotion pattern. When a person peak frequency in alpha band is relatively high, it is indicated that a person is in an unpleasant emotion.

All the features parameters of speech signal from male and female has been obtained by using autocorrelation and linear predication method in order to obtain mean fundamental frequency, standard deviation fundamental frequency, maximum fundamental frequency, minimum fundamental frequency, mean intensity, first formant frequency and pause length ratio. Among these features parameters, mean fundamental frequency, standard deviation fundamental frequency and mean intensity show a good result in order to differentiate between angry emotion and no angry emotion. This is because, during high degree emotion (angry), speech was uttered faster with bigger volume and raised the tone due to produced changes in respiration and an increase in muscle tension, which increase the vibration of the vocal folds, and variation of fundamental frequency also increases. Therefore, mean pitch, standard deviation pitch and mean intensity was much higher. However, during in low degree emotion (neutral and sadness), there were no much alteration in the fundamental frequency and the value of fundamental frequency was very low as well as the intensity was much lower compared to high degree emotion.

In addition to the EEG and speech analysis, the heartbeat measurement is done to check the level changes while people having different emotion. This is sensed by using a high intensity type Light Emitting Diode (LED0 and LDR. The finger is placed between the LED and LDR. The skin may be illuminated with visible light (red) using transmitted or reflected light for detection. The very small changes in reflectivity or in transmittance caused by the varying blood content of human tissue are almost invisible. Its characteristic is related to human emotion and changes the body condition directly. From the experiment result, it is proved that human heart will beat faster than normal when they are angry.

ACKNOWLEDGMENT

This work has been supported financially by the Office for Research, Innovation, Commercialization and Consultancy Management (ORICC), Universiti Tun Hussein Onn Malaysia and the facilities and information are provided by Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia. Authors would like to thank A. Savran, et al. for sharing their EEG signals database as an open source. Authors also have to thank Sir German Gomez-Herrero for sharing a useful AAR toolbox.

REFERENCES

[1] Plutchik, R., *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion.* New York: Academic, 1980.
[2] Ekman, P. , *Basic Emotions.* Sussex, U.K.: John Wiley & Sons, Ltd, 1999.
[3] Russell, J.A., "Affective Space is Bipolar". *Journal of Personality and Social Psychology*, vol. 37, no. 3, 1979, 345–356.
[4] Wang Yu, "Research and Implementation of Emotional Feature Classification and Recognition in Speech Signal," *IITAW'08*, 2008, pp. 471-474.
[5] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J. Taylor, "Emotion Recognition in Human-Computer Interaction," *IEEE Signal Process. Mag*, vol.18, Jan 2001, pp. 32-80.
[6] Takahiro Sunaga and HirooIkehira, "Measurement of the Electrical Properties of Human Skin and the Variation among Subjects with Certain Skin Conditions," *Physics in Medicine and Biology*, vol. 47, no. 1, 2002.
[7] Alberto de Santos Sierra, Carmen Sánchez Ávila, Javier Guerra Casanova, and Gonzalo Bailador del Pozo, "A Stress-Detection System Based on Physiological Signals and Fuzzy Logic," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 10, 2011, pp. 4857–4865.
[8] Al-Ani, A. & Al-Sukker, A., "Effect of Feature and Channel Selection on EEG Classification," *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE* , 2006, pp. 2171 - 2174 .
[9] Hagemann, D., Naumann, E. & Thayer, J.F., "The Quest for the EEG Reference Revisited: A Glance from Brain Asymmetry Research'".*Psychophysiology , 38*, 2001, pp. 847-857.
[10] Hoffmann, U., Vesin, J.-M., Ebrahimi, T. & Diserens K., "An Efficient P300-Based Brain Computer Interface for Disabled Subjects," *Journal of Neuroscience Methods* , vol. 167, no. 1, 2008, pp. 115-125.
[11] Horlings, R., Datcu, D. & Rothkrantz, L.J.M., "Emotion recognition using Brain Activity," *In Proceedings of CompSysTech'2008*, Gabrovo, Bulgaria, 2008.
[12] Murugappan, M., Nagarajan, R. & Yaacob, S., *Discrete Wavelet Transform Based Selection of Salient EEG Frequency Band for Assessing Human Emotions.* Seriab, Kangar, Perlis: InTech., 2010.
[13] A. Savran, et al., *eNTERFACE*., 2006, from eNTERFACE'06: www.enterface.net
[14] Deller, J. R, Hansen, J. H. L. and Proakis, J., *Discrete-Time Processing of Speech Signals.* Wiley-IEEE pres. 2000.
[15] Keikichi Hirose, Hiroya Fujisaki and Shigenobu Seo, "A Scheme for Pitch Extraction of Speech Using Autocorrelation Function with
[16] Length Proportional to the Time Lag," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1992, pp. 149-152.
[17] M. Morales-Perez, J. Echeverry-Correa, A. Orozco-Gutierrez and G. Castellanos-Domingue, "Feature Extraction of Speech Signal in Emotion Identification" *EMBS*, 2008, pp.20-24.