International Journal of Soft Computing 11 (6): 437-443, 2016
ISSN: 1816-9503

A Study on Emotional Identification Using Facial Electromyogram Signals and Neural Networks

G. Charlyn Pushpa Latha and M. Mohana Priya
Faculty of Engineering, Karpagam University, Coimbatore, India

Abstract: This study attempts to state that statistical signal processing treats signals as stochastic processes. It deals with the statistical properties to process signals and extract significant features. Being a versatile feature extraction method, it is also used in different areas such as natural language processing, bio-signal processing and sonar. In this research, it has been examined that the Facial Electromyography signals (FEMG) are processed by applying the statistical features in order to extract features for categorizing six emotions namely, happy, fear, neutral, sad, disgust and anger. Twenty subjects have taken part in this experimental study. The statistical features namely, kurtosis, skewness, moment, range, median absolute deviation and mean have been used to derive the significant features. Six emotions have been identified by applying the statistical features as input to neural network models. There are four neural network models namely, Cascade network, Elman network, Layered recurrent network and feed forward network have been used and compared to identify an efficient network for emotional identification. The performances of the networks in identifying the six emotions were in the range of 87.56-98.33%.

Key words: Facial electromyography, statistical features, elman neural network, feed forward neural network, cascade neural network, layered recurrent neural network

INTRODUCTION

Emotions are innate and discrete to specific event or condition. Emotional experience varies from person to person in same situations at variable instances of time. Emotional collaboration is the foremost quantity which deals with social communications in persons. It explicitly lies on facial muscle movements. It has also been defined as a critical state that comparatively relates to a specific occurrence (Westerink *et al*., 2008). Studies on emotions were improved considerably over the years on various arenas, subsidizing to different fields comprising, medicine, history, neuroscience, endocrinology, sociology etc. The functioning emotions are principally termed as arguments in bi-coordinates emotional area. It consists of emotional arousal factor and valence factor. Valence factor indicates the overall strength of emotional capabilities ranging between negative and positive, whereas arousal factor denotes emotional strength between calm emotion and excited emotion (Lang, 1995). Various possibilities have been discovered by the researchers to endow emotional skills in computers for Human Computer Interaction (HCI). Hence emotional recognition system could be entailed in existing environment which is trustworthy, perfect, simple and pliable by nature.

FEMG is an electromyography technique that measures the activity of facial muscles by detecting and amplifying the tiny electrical impulses and they are generated by the muscle fibres when they are in contract. FEMG primarily focuses on two major muscle groups in the face, the corrugator supercilii group is associated with frowning and the zygomaticus major muscle group which is associated with smiling. Inorder to detect the facial emotions accurately, scientists endeavour the accomplishment of a purpose on FEMG (Chen and Huang, 2000; Chen and Chung, 2004). Facial expressions are comparatively efficient with other nonverbal communications like body posture. Facial expressions seem to be highly inspiring sensations that focuses to sentient (Zuckerman *et al*., 1986). Facial lexes and speech systems based on emotional recognition were also described in the earlier studies. Researchers around the world have developed several techniques related to sensors such as Electro Encephalo Gram (EEG), heart rate, skin conductance, respiration rate and EMG signals (Wagner *et al*., 2005; Wang and Guan, 2006). FEMG helps to identify the emotional reactions from the face when compared to other physiological signals. Six different emotional conditions namely, happy, anger, fear, disgust, sad and neutral have been used in this study. Statistical parameters namely, median absolute deviation, mean, range, skewness, moment and kurtosis parameters and

Corresponding Author: G. Charlyn Pushpa Latha, Faculty of Engineering, Karpagam University, Coimbatore, India

neural networks namely, Cascade neural network, Feed forward neural network, Elman neural network and layered recurrent neural network are used. The investigation supports the related work which is illustrated in Part 2. Part 3 examines the FEMG data collection techniques, pre-processing methods and feature extraction technique. Part 4 demonstrates the experimental outcomes. The discussions on the methods with the certain inferences are illustrated at the end of this research.

Previous work: Emotions are highly independent and explicit to an individual occurrence or state. All individuals do not experience the same emotional strength to similar situations. Experiments deal with EMG techniques which was started in 1666. Rigas *et al.* (2007) used the classification algorithms namely, Random forest classifiers and K-Nearest Neighbour (K-NN) for happiness, disgust and fear. They designed an emotion recognition system and obtained the recognition rate of 69% for these emotions. Wong *et al.* (2010) used the emotional features namely, sadness, anger, pleasure and joy, derived the highest recognition rate of 89.25% with the particle swarm optimization of synergetic neural classifier. Yang and Yang (2011) derived the recognition rates of 91.67% for happiness and disgust by using SVM and back propagation network. A fuzzy C-means classifier is used to emotions namely, rest, smile, frown, rage and pulling up eyebrows was proposed by Hamedi *et al.* (2011). Principal component analysis and higher order statistics were adopted by Jerritta to design a technique for categorizing the emotions from the facial EMG signals. Kim and Andre (2008) identified 4 emotional conditions namely, anger, sad, pleasure and joy achieved 95% by using extended LDA classifier.

Gibert *et al.* (2009) analyzed the FEMG signal with a sampling frequency of 25Hz and obtained the mean recognition rate of 92% for six emotions. Paul *et al.* (2014) analyzed the FEMG signals with a sampling frequency of 250 Hz. Classification and regression tree, Linear discriminant analysis, self organizing, ap and naïve bayes classifiers were used for negative emotions like sadness, fear, surprise and stress by Cohen *et al.* (2003) by using a sampling frequency of 30 Hz. Shih *et al.* (2013) used a sampling frequency of 100 Hz to analyze the FEMG signals. They used 3 databases namely, Cohn-kanade database, MMI database and lab078 database obtaining an average accuracy of 89%. Literature review reveals that very few studies on neural networks have been experimented on identifying emotions from FEMG signals. This study proposes to use for both static and dynamic networks to provide a novel approach in emotional identification.

MATERIALS AND METHOD

Femg signal acquisition: Emotions can be stimulated either through visual, audio-visual and recalling former emotional incidents or with audio clips Y (Daabaj, 2002; Hongo *et al.*, 2000). The audio-visual method is used for inducing emotions in this experimental study. FEMG data are recorded with the help of an analog digital instrument bio-amplifier. Faces of subjects are instructed to remove makeup for better signal detection, five gold plated electrodes are used and placed on the locations such as depressor anguli oris, Corrugator supercilli, Levator muscle labii superioris, Orbicularis oris muscle and at the reference point as shown in Fig. 1. The facial muscles are chosen to detect activity from the regions as described below (Massaro *et al.*, 2000):

- Corrugator supercilli muscle is chosen which moves from the eyebrow to the middle end and down
- Levator labii superioris muscle shrinks the nose, enlarges nasal divisions and uplifts the upper lip
- Orbicularis oris muscle restricts the skin around the eye
- Depressor anguli oris muscle controls over the size and shape of the mouth openings

The sampling frequency of the FEMG signals is chosen as 200 Hz. Audio-visual clippings from the movies modelled by Tomarken *et al.* (1990) have been selected to invoke the six emotions. Eleven males and nine females in the age group of 18-50 years volunteered as subjects for the experiments. FEMG signals are recorded for all the six emotions namely, happiness, anger, disgust, fear, sadness and neutral. Separate audio-visual clips are used for each emotion. Ten trials are recorded for each emotion per session. Data from two sessions are collected. Subjects are given a break for 20 min in-between the recordings. A notch filter is applied during the recording process to remove power artifacts.

Spectral analysis estimation: A spectral analysis using short time fourier transform is conducted to reveal the frequency of the signals for the six emotions as shown in Fig. 2 the spectrum shows that all emotions have a frequency range below 10 Hz. Using chebyshev filter the signals are bandpass filtered with a band length of 1 Hz. Ten segments are used with bands ranging from 0.1-10 Hz.

Statistical feature extraction: Six statistical algorithms are applied to derive the statistical features which are detailed as:

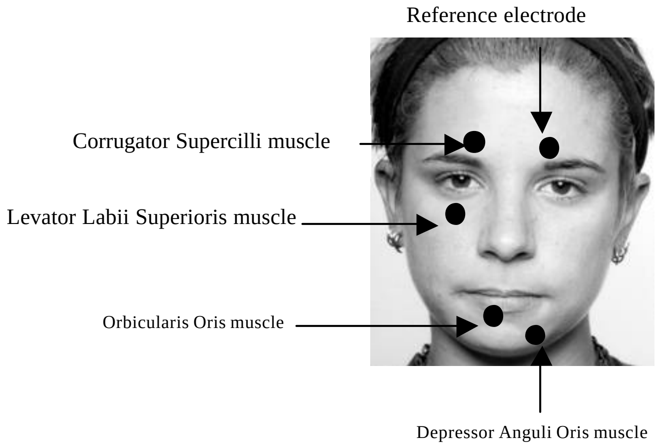

Fig. 1: Positions of electrodes in FEMG

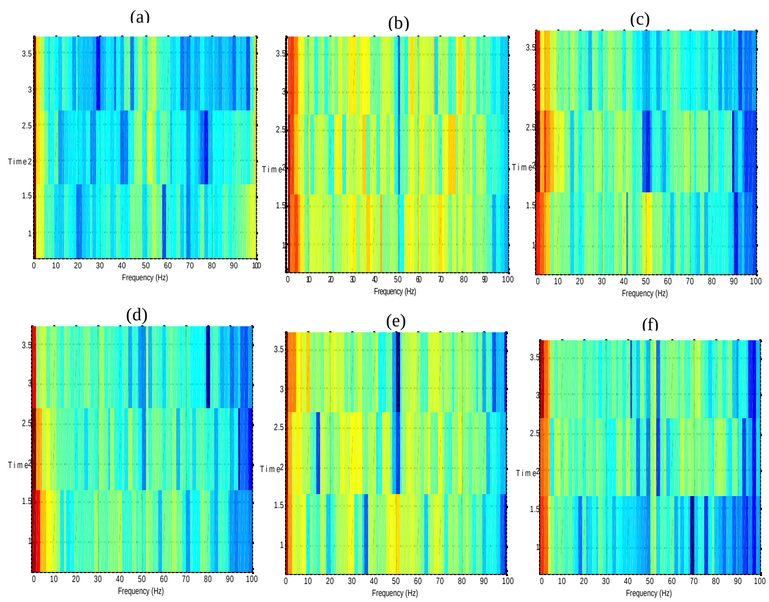

Fig. 2: Spectral analysis of the six emotions for subject 4: a Anger; b) disgust; c) fear; d) happy; e) neutral and f) sad

Mean: This statistical parameter computes the average rate of the FEMG signal. The formula used for the estimation of mean:

$$\mu_m = 1/T\sum_{b=1}^{T} Z_b \tag{1}$$

where, Z_b b = 1, 2, 3, ..., T is the raw FEMG data; T is denotes size of raw FEMG data.

Median absolute deviation : This parameter evaluates the mean absolute deviation value of the signal which is the average value of total variations of each data in signal. It sets along with the average of all the data containing in that particular signal. Extraction of the median absolute deviation is done either using median or mean. The reason of median is opted is for the summation of group deviation when observed from median is least on ignoring the signs. The equation used for calculation:

$$\frac{\sum |Z_b - MD|}{T} \tag{2}$$

where, T is indicates size of the raw FEMG data Z_b = 1, 2, 3, ..., T is the raw FEMG data indicates the median of the signal.

Range: The sign variation of the highest and lowest data in the signal set is referred as the signal range.

Moment: Principal moment of order k of a signal is:

$$x_k = E(b-\mu)^u \quad (3)$$

where, E(b) is the estimated value of b which is the dimension of the mid probability function.

Skewness: The standardized forms of third and fourth term are skewness and kurtosis .The asymmetry degree of a circulation is about their average is symbolized by Skewness which describes the distribution shape. It represents also the third central moment of X, divided by the cube of its standard deviation. The formula for calculation:

$$\text{Skewness}(\upsilon_3\hat{}) = \sum_{b=1}^{T} \frac{(Z_b-\mu_m)^3}{(T-1)\sigma^3} \quad (4)$$

Where:

μ_m = The average

σ_m = The standard deviation of the pre-processed signal Z_b

T = Represents the size of the initial value of FEMG signal

Kurtosis: The degree of comparative immensity at the end of dissemination in connection to the normal dissemination is referred to as kurtosis. It also represents the fourth central moment of X, divided by fourth power of its standard deviation. The equation used:

$$\text{Kurtosis}(\upsilon_4\hat{}) = \sum_{n=1}^{N} \frac{(Z_b-\mu_m)^4}{(T-1)\sigma_m^{\ 4}} - 3 \quad (5)$$

Algorithm for Statistical Feature Extraction

Set the Initial Parameters as follows:

1. Set the window size = ½ * (sampling frequency) as per Nyquist Criteria.
2. Apply Chebyshev filter to the frequency bands.
3. Restrict the number of bands to 10 based on the spectrum analysis.

Input:

1. For each emotion, (h =2000 data samples) are given as input to the feature extraction algorithm
2. Find the transpose (ht= h') of the input signal.
3. Calculate the rows and columns as [m1 n1] = size (ht) of each input data file.

Process:

1. Calculate the median absolute deviation: (mad (y1)) of the input samples.
2. Calculate the range: range (y1) of the input samples.
3. Calculate the mean: mean (y1) of the input samples.
4. Calculate the moment: moment (y1, 10) of the input samples.
5. Calculate the skew: skewness (y1) of the input samples
6. Calculate the kurt: kurtosis (y1) of the input samples.
7. Repeat the process until the row value of the data samples are met

Output:

1. 120 features are extracted as output comprising of 20 values for each of the statistical features namely, median absolute deviation, range, mean, moment, skewness and kurtosis for the six emotions namely anger, disgust, fear, happy, neutral and sad for all the 20 subjects.

Neural network based emotional identification: One static and three dynamic neural networks are compared to an efficient neural network algorithm for emotional identification. The static Feed Forward Neural Network (FFNN) is a robust model that was developed in the earlier times and there is no feedback among the individual layers. FFNN applies a supervised learning technique and is also known as back propagation networks. Back propagation algorithm involves in two phases namely, the forward and the backward phases. In the forward phase, the network parameters are fixed and the input signal is propagated through the network layer and it also computes the error signal. In the backward phase, the error signal is propagated through the network in backward direction and the adjustments are applied to the parameters of the network so as to reduce the error. The training function used in FFNN algorithm is 'trainlm' which is a network training function that updates weight and bias states according to levenberg-marquardt optimization.

The dynamic Elman Neural Network (ENN) comprises with four layers namely, input layer, output layer, hidden layer and context layer (Hema *et al.*, 2008). It is also called as feedback neural network because it has extra feedback connection from the output hidden layer to input through context layer. It adds the ability to learn the temporal characteristics of the data set. The context layer in this network stores the previous state of the hidden layer which helps in improving the classification rate and the training time of the network when compared to feed forward networks.

Cascade Forward Network (CFNN) includes weight connection from the input layer to each layer and from each layer to the consecutive layers. It also uses the back propagation algorithm for updating weights and each layer of neurons relates to all the previous layer of neurons (Hema, 2010). The training function used in CFNN algorithm is also 'trainlm'.

An improved version of elman network is the Layered Recurrent Neural Network (LRNN). Each and every layer in this network has a feedback loop, with single delay except the last layer. LRNN uses bayesian regulation of back propagation algorithm and it has an arbitrary number of layers to have arbitrary transfer functions in each layer. (Paulraj *et al.*, 2009; Hema *et al.*, 2012; Hema *et al.*, 2009).

Data from twenty subjects are analyzed and networks are developed for each subject data and 80 networks are modeled to investigate the best network model for emotional identification. A total of 120 features are given as input to each network. All networks are designed using 120 input neurons, 3 output neurons and 15 hidden neurons. The hidden neurons are chosen by trial and error process. 75% of the data set is used for training and 100% of the data set is used for testing the networks. The testing and training error tolerance are fixed as 0.05 and 0.0001 respectively.

RESULTS AND DISCUSSION

The classification performance of the network model chosen in this study is illustrated graphically in Fig. 3. Comparing all the network models, LRN achieved the highest classification accuracy of 98.33% for subject 4. The LRN model has recurrent connection with a tap delay associated which allows the network to have an infinite dynamic response to time series input data. It also uses a backpropogation learning algorithm which uses gradient descent method to update the weights in order to minimize the loss function.

Out of the 20 subjects who participated in the experiment, eleven were males and nine were females. Comparing the overall performances of all the 80 network models, the highest accuracy was achieved by CFNN network for all subjects except subject 3 and the recognition accuracy varied from 80.83-97.58%. It was also observed that among the 20 subject data, the data of subject 4 has the highest classification performance at 98.33% (LRN), 97.58% (CFNN) and 94.5% (FFNN).

From the empirical results, it is observed that the study among the four network models using statistical features, the CFNN network has better recognition rates for the six emotional states. It is also observed that male subjects had comparatively better recognition rates in comparison to females. The age of the subjects did not play a vital role in improving the recognition rates. The experimental results validate the feasibility of recognizing emotions of individual using FEMG signals and neural networks. In previous research, studies on FEMG using the six emotions namely, anger, surprise, disgust, happiness, sadness and neutral with gaussian model classifier, the highest classification accuracy obtained is 92% (Gibert *et al*. 2009). For EEG studies using six emotions namely satisfied, surprise, happy, sad, fear and angry and with the SVM classifier the highest classification accuracy obtained is 75.17% (Latha *et al*., 2013). Hence, it is observed that the results obtained in the study are comparatively higher than the other studies conducted on six emotional states.

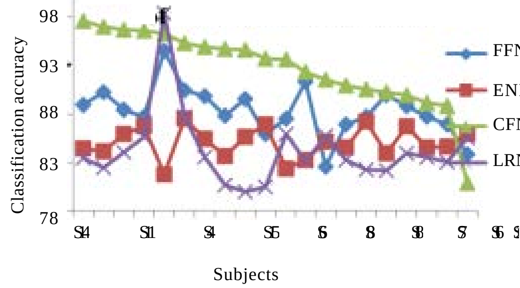

Fig. 3: Comparison of classification accuracies of the four network model for 20 subjects

CONCLUSION

Recognizing emotions by F EMG signals have drawn recent attention. FEMG analysis is a well-thought and delicate method for analysing the individual's emotional condition. Anyhow, it has drawbacks for some real life claims due to impact of the parameters affecting the facial muscle movements. Eleven males and nine females participated in the experiment. In this study, the statistical parameters are analysed to classify the six emotions disgust, fear and anger, neutral, sadness and happy using four neural network models namely CFNN, LRN, ENN and FFNN. The classification accuracies obtained for four network models namely FFNN, ENN, CNN and LRN are 94.5, 87.56, 97.58 and 98.33%, respectively. In this study, out of the four network models, the CFNN model has the highest performance in emotional recognition. Performances of data from male subjects have been observed to have better recognition rates when comparing to females. Emotions namely disgust, happy, neutral and sad were observed well from the subjects whereas the subjects lack training in anger and fear emotions. Out of the six emotions proposed in this study, disgust emotion was the best exhibited emotion. The recognition rates achieved in this study are relatively higher in comparing with similar studies (Latha *et al*., 2013). However, the performances of the proposed algorithms are to be verified on online FEMG signals to develop a real time FEMG based emotion recognition system.

REFERENCES

Chen, C.N.H.C.H. and H.Y. Chung, 2004. The review of applications and measurements in facial electromyography. J. Med. Biol. Eng., 25: 15-20.

Chen, L.S. and T.S. Huang, 2000. Emotional expressions in audiovisual human computer interaction. Proceeding of the 2000 IEEE International Conference on Multimedia and Expo, 2000 ICME 2000, July 30-August 2, 2000, IEEE, Urbana, Illinois, ISBN:0-7803-6536-4, pp: 423-426.

Cohen, I., N. Sebe, A. Garg, L.S. Chen and T.S. Huang, 2003. Facial expression recognition from video sequences: Temporal and static modeling. Comput. Vision Image Understanding, 91: 160-187.

Daabaj, Y., 2002. An evaluation of the usability of human-computer interaction methods in support of the development of interactive systems. Proceedings of the IEEE International Conference on System Sciences, Jan. 7-10, IEEE Xplore Press, Hawaii, pp: 1830-1839.

Gibert, G., M. Pruzinec, T. Schultz and C. Stevens, 2009. Enhancement of human computer interaction with facial electromyographic sensors. Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7, November 23-27, 2009, ACM, New York, USA., ISBN:978-1-60558-854-4, pp: 421-424.

Hamedi, M., I.M. Rezazadeh and M. Firoozabadi, 2011. Facial gesture recognition using two-channel bio-sensors configuration and fuzzy classifier: A pilot study. Proceedings of the 2011 International Conference on Electrical, Control and Computer Engineering (INECCE), June 21-22, 2011, IEEE, Tehran, Iran, ISBN:978-1-61284-229-5, pp: 338-343.

Hema, C.R., 2010. Recognizing motor imagery using dynamic cascade feed-forward neural networks. Proceeding of the 2010 6th International Colloquium on Signal Processing and Its Applications (CSPA), May 21-23, 2010, IEEE, Perlis, Malaysia, ISBN:978-1-4244-7121-8, pp: 1-3.

Hema, C.R., M.P. Paulraj and A.H. Adom, 2012. Improving classification of EEG signals for a four-state brain machine interface. Proceedings of the 2012 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), December, 17-19, 2012, IEEE, Coimbatore, India, ISBN:978-1-4673-1664-4, pp: 615-620.

Hema, C.R., M.P. Paulraj, R. Nagarajan and Y.A.H. Sazali, 2008. Brain machine interface: Analysis of segmented eeg signal classification using short-time PCA and recurrent neural networks. Int. J. Electr. Electron. Eng. Iraq, 4: 77-85.

Hongo, H., M. Ohya, M. Yasumoto and K. Yamamoto, 2000. Face and hand gesture recognition for human-computer interaction. Proceedings of the 15th IEEE International Conference on Pattern Recognition, Sept. 3-8, IEEE Xplore Press, USA., pp: 921-924.

Kim, J. and E. Andre, 2008. Emotion recognition based on physiological changes in music listening. IEEE Trans. Pattern Anal. Mach. Intell., 30: 2067-2083.

Lang, P.J., 1995. The emotion probe: Studies of motivation and attention. Am. Psychol., 50: 372-385.

Latha, G.C.P., C.R. Hema and M.P. Paulraji, 2013. Neural network based classification of human emotions using Electromyogram signals. Proceedings of the 2013 International Conference on Advanced Computing and Communication Systems (ICACCS), December 19-21, 2013, IEEE, Coimbatore, India, ISBN:978-1-4799-3507-9, pp: 1-4.

Massaro, D.W., 2000. Perceptual interfaces in human computer interaction. Proceedings of the IEEE International Conference on Multimedia Expo, July 30-Aug. 2, IEEE Xplore Press, New York, pp: 563-566.

Paul, G.M., F. Cao, R. Torah, K. Yang and S. Beeby *et al*., 2014. A smart textile based facial EMG and EOG computer interface. IEEE. Sens. J., 14: 393-400.

Paulraj, M.P., R.B. Ahmad, C.R. Hema and F. Hashim, 2009. Estimation of mobile robot orientation using neural networks. Proceedings of the 5th International Colloquium on Signal Processing its Applications, 2009, CSPA 2009, March 6-8, 2009, IEEE, March 6-8, 2009, ISBN:978-1-4244-4151-8, pp: 42-46.

Rigas, G., C.D. Katsis, G. Ganiatsas and D.I. Fotiadis, 2007. A User Independent, Biosignal Based, Emotion Recognition Method. In: User Modeling, Cristina, C., M. Kathleen and P. Georgios (Eds.). Springer, Berlin, Germany, ISBN:978-3-540-73077-4, pp: 314-318.

Shih, C.H., C.H. Ting and L.H. Chung, 2013. Video based facial expression recognition using hough forest. Proceeding of the 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), October 29-November 1, 2013, IEEE, Hsinchu, Taiwan, ISBN:978-986-90006-0-4, pp: 1-9.

Tomarken, A.J., R.J. Davidson and J.B. Henriques, 1990. Resting frontal brain asymmetry predicts affective responses to films. J. Personality Social Psychol., 59: 791-801.

Wagner, J., J. Kim and E. Andre, 2005. From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification. Proceeding of the 2005 IEEE International Conference on Multimedia and Expo, July 6, 2005, IEEE, Augsburg, Germany, ISBN:0-7803-9331-7, pp: 940-943.

Wang, Y. and L. Guan, 2006. Recognizing human emotion from the audio visual information. Proceeding of the International Conference on Acoustics, Speech and Signal Processing, March 18-23, 2006, IEEE, Philadelphia, Pennsylvania, pp: 1125-1128.

Westerink, J.H., D.B.E.L. Van, M.H. Schut, H.J. Van and K. Tuinenbreijer, 2008. Computing Emotion Awareness Through Galvanic Skin Response and Facial Electromyography. In: Probing Experience, Joyce, H.D.M.W., O. Martin, J.M.O. Therese, W.F. Pasveer and D.R. Boris (Eds.). Springer, Netherlands, Europe, ISBN:978-1-4020-6592-7, pp: 149-162.

Wong, W.M., A.W. Tan, C.K. Loo and W.S. Liew, 2010. PSO optimization of synergetic neural classifier for multichannel emotion recognition. In: Proceeding of the 2010 Second World Congress on Nature and Biologically Inspired Computing (NaBIC), December 15-17, 2010, IEEE, Malacca, Malaysia, ISBN:978-1-4244-7377-9, pp: 316-321.

Yang, S. and G. Yang, 2011. Emotion recognition of EMG based on improved L-M BP neural network and SVM. J. Software, 6: 1529-1536.

Zuckerman, M., B.M. DePaulo and R. Rosenthal, 1986. Humans as Deceivers and Lie Detectors. In: Nonverbal Communication in the Clinical Context, Blanck, P.D., R. Buck and R. Rosenthal (Eds.). Pennsylvania State University, Pennsylvania, USA., pp: 13-35.