

The 9th International Conference on Cognitive Science

EEG emotion recognition based on the dimensional models of emotions

Marini Othman^{a,*}, Abdul Wahab^a, Izzah Karim^a, Mariam Adawiah Dzulkifli^b, Imad Fakhri Taha Alshaikli^a

^a*Department of Computer Science, Kulliyah of Information and Communication Technology, International Islamic University Malaysia, Jalan Gombak, Kuala Lumpur 53100, Malaysia*

^b*Department of Psychology, Kulliyah of Islamic Revealed Knowledge and Human Sciences, International Islamic University Malaysia, Jalan Gombak, Kuala Lumpur 53100, Malaysia*

Abstract

In this paper, we propose a method for EEG emotion recognition which is tested based on 2 dimensional models of emotions, (1) the rSASM, and (2) the 12-PAC model. EEG data were collected from 5 preschoolers aged 5 years old while watching emotional faces from the Radboud Faces Database (RafD). Features were extracted using KSDE and MFCC and classified using MLP. Results show that EEG emotion recognition using the 12-PAC model gives the highest accuracy for both feature extraction methods. Results indicated that the accuracy of EEG emotion recognition is increased with the precision of the dimensional models.

© 2013 The Authors. Published by Elsevier Ltd.

Selection and/or peer-review under responsibility of the Universiti Malaysia Sarawak.

Keywords: brain signals; valence-arousal model; preschoolers; children; emotions; machine learning; classification.

1. Introduction

Researches in the area of brain computer interfaces (BCI) have provided evidences that human emotion can be quantified using EEG emotion recognition systems [1][2]. These studies, however, have mainly focused on distinguishing separable basic emotion widely known among psychologists as the discrete emotion model. At the moment there is no consensus on the number of basic emotions. However, most researchers have agreed to accept six emotions namely happy, sad, fear, anger, disgust and surprise [3]. The discrete emotion model has been mainly criticized for its incapability for capturing human emotions in relation to other emotions [4]. Nevertheless, this model has been widely accepted due to its simplicity, high plausibility and interpretability [5].

Another popular approach for distinguishing emotions is by placing them in a few fundamental dimensions based on human appraisal for specific emotional events [6], which is perceived as giving fully complementary description of the emotion [4]. For example, a two-dimensional circumplex model of affect was proposed in [7] to quantify an emotion based on the level of valence and the intensity of arousal. Some of the early circumplex models have

* Corresponding author. Tel.: +603-61965601 ; fax: +603-61965179 .

E-mail address: omarini@iiu.edu.my

simply placed affective states on specific affective regions leaving many areas largely uncharted. Recent model such as in [8] allows the existence of coordinate structure of emotions providing opportunities for higher precision of emotion recognition.

The recognition of human emotions is normally treated as a pattern classification problem with accurate classification can be achieved by identifying appropriate feature spaces and designing a classifier that closely model the classification problem based on the selected feature space [9]. Thus, the objective of this paper is to propose a method for EEG emotion recognition based on 2 dimensional models of emotions, namely the rSASM [5] and the 12-PAC model [7]. In our work, KSDE and MFCC are employed as the feature extraction method while MLP is used as the classifier. The location of specific emotions in the dimensional models, at times termed as the emotion primitive values [5] may serve as targets of the classifier.

Another noticeable gap in the current emotion recognition studies is that most of these researches have concentrated on the psycho-physiological investigation of adults' emotion while the developmental years have been largely ignored. One of the reasons might be due to the lack of developmental data for analysis, since challenges in collecting children's data are magnified compared to working with adults [10]. Evidences regarding children's conception of emotion have also been perceived as a puzzle [11]. While some researchers believe that children's acquisition of emotion can be as early and as fast as 2 years old [12], others have indicated that emotion perceptions are acquired gradually due to the different stages of children's brain maturity [11][13].

To fill in this gap, our analysis is based on the EEG data during developmental years; specifically preschoolers aged 5 years old. The EEG feature extracted data shall be classified using 2 approaches, (1) homogenous classification for reflecting the diverse stages of children's brain maturity, and (2) heterogeneous classification for generalization of children's emotion due to the assumption of 'fast acquisition' of emotion. In this case, a homogenous classifier refers to a neural network that learns and constructs classifier for each subject for capturing individual differences, rather than trained for the whole data in global network training (i.e. heterogeneous classification). Results from this study may leads towards the understanding of children emotions for the purpose of intervention in assisting brain developmental disorder such as ADHD and autism.

2. The Dimensional Models of Emotions

2.1. The recalibrated Speech Affective Space Model (rSASM)

The recalibrated Speech Affective Space Model (rSASM) is previously proposed for recognizing cultural-influenced speech emotion [4]. In the rSASM, the emotion primitive values are derived from the 4 quadrant of emotions which serves as the intended output for feature classification. In Quadrant II, however, emotion angry in the original rSASM is replaced with fear since emotion fear is widely investigated in the area of cognitive neuroscience compared to angry. The interpretation of the rSASM for our work is displayed below:-

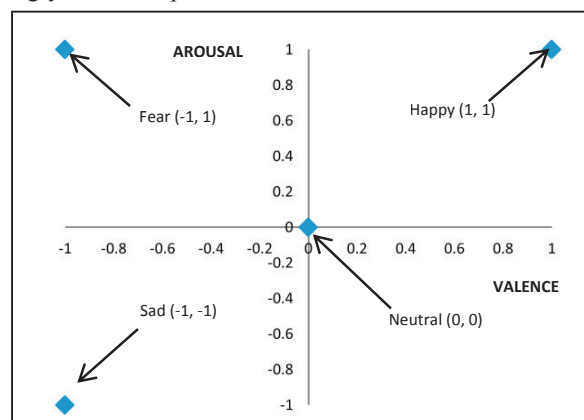


Fig.1. The intended output for different emotions based on the rSASM [2]

2.2. The 12-Point Affective Circumplex (12-PAC)

The recently proposed 12-Point Affective Circumplex (12-PAC) integrates different dimensional model of moods and emotions based on 4 correlational studies, hence producing a geometrical model that allows for higher precision in emotional profiling. To the best of our knowledge, the 12-PAC model has not been tested with any EEG based emotion recognition system.

In this model, the emotion primitive values are based on Yik et al's [7] placement of external variables using the CIRCUM-extension method. Figure 2 below shows the exact location of emotion fear, happy (described as pleasant in [7]) and sad. Neutral was not stated in the 12-PAC model but interpreted as [0, 0] as in rSASM. In the 12-PAC model, $\theta+$ is defined as the estimated angle of each variable while $\zeta+$ is defined as the square root of the proportion of the external variable explained by the CIRCUM model for the 12-PAC structure (i.e. magnitude of the relation).

Using both of the $\theta+$ and $\zeta+$ values described in Table 6 from [7], the exact emotion primitive values of valence, v_{pc} and arousal, a_{pc} for the 12-PAC model can be determined as follows:-

$$v_{pc} = \zeta+ \cos \theta+ \quad (1)$$

$$a_{pc} = \zeta+ \sin \theta+ \quad (2)$$

Figure 2 below shows the intended output based on the 12-PAC model:-

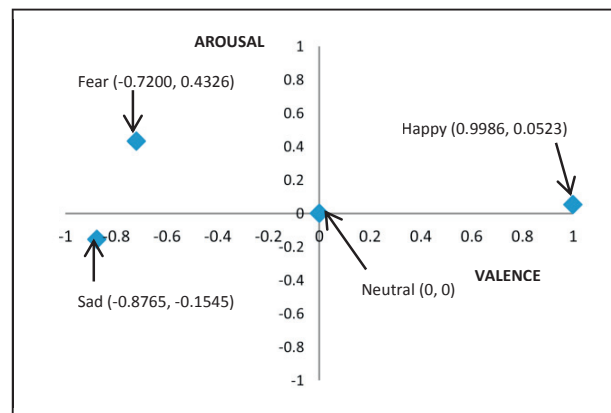


Fig. 2. The intended output for different emotions based on the 12-PAC model [7]

2.3. Output encoding schemes for rSASM and 12-PAC

Drawing from the dimensional models discussed in Section 2.1 and 2.2 above, the final output encoding schemes for the rSASM and 12-PAC can be summarized as follows:-

Table 1. Summary of the output encoding schemes for rSASM and 12-PAC

Emotion	rSASM	12-PAC
Happy	{1,1}	{0.9986, 0.0523}
Fear	{-1,1}	{-0.7200, 0.4326}
Sad	{-1,-1}	{-0.8765, -0.1545}
Neutral	{0,0}	{0,0}

3. The EEG Emotion Recognition System

3.1. Data collection and pre-processing

The sample of our study is 5 preschoolers aged 5 years old. All parents of the participants gave informed, signed consent. EEG data were collected in a lighted, temperature controlled room. All children were instructed to close their eyes for 1 minute and open their eyes for another minute for baseline recording. The children watched emotionally related human facial expression from the Radboud Faces Database (RafD) [14] for the affective state of happy, sad, neutral and fear. Each sets of pictures last 1 minute.

Brain signals are collected over the frontal lobes (F3, F4) with the sampling frequency of 250Hz while Cz was used as the reference lead. The placement of EEG probes is based on the understanding of activity in the respective brain region, whereby past research has indicated that the frontal lobes are responsible for emotional regulations [15]. Emotional face processing is also known to evoke ERP responses in the frontal regions [16]. Impedances were kept below 10k Ω throughout the experiments.

Consistent with the work of [17], a minimum of 40 seconds emotion data is analyzed in this study. Prior analysis using Cronbach's alpha on the resting state EEG indicated data reliability, with all of the obtained reliability coefficients are above 0.85. Elliptic filters were applied on the signals for retaining the theta (4-8Hz) and alpha band (8-13Hz) that correlates with emotional experiences [18].

3.2. Features Extraction

The Mel-Frequency Cepstral Coefficients (MFCC) and Kernel Density Estimation (KDE) were employed in this study as the feature extraction methods. The usage of MFCC is based on the behavior of the mel-frequency that follows a linear spacing below 1 kHz and a logarithmic spacing above 1 kHz. In our case with the sampling frequency of 250Hz, a linearity assumption can be used on our EEG data. Twelve linearly spaced mel filterbanks were computed for each frequency band and EEG channels of F3 and F4 for obtaining the mel cepstrum. This resulted in 24 features for each channel which are later combined, forming a 48-features matrix.

Kernel density estimation (KDE) is a non-parametric approach that computes the probability distribution function of random variables without having to assume the distributions of data samples [19]. In this approach, the kernel estimators smooth out the contribution of each experimental data point over a local area which can be used for feature extraction. Our analysis is based on the density estimate evaluated at 10 equally spaced feature points that covers the range of data in each frequency band and the collected EEG channels. Thus, a final 40-feature matrix was obtained for this feature extraction method.

3.3. Classification

The Multi-Layer Perceptron (MLP) is quite popular among researchers for its learning capability, and has been proven to be viable for EEG emotion recognition systems [2][19][20]. The MLP is a feed-forward artificial neural network with at least 1 hidden layer of neurons. The neurons are responsible for computing the weighted sum of the incoming signals, to yield net inputs, and pass these values through an activation function to yield the neurons' activation values.

Our MLP network architecture is set to 2 layers with 10 neurons in each layer. Training goal is set to 0.01. For the homogenous classification approach, the training iterations are fixed at 10,000 epochs while training iterations for heterogeneous classification are varied from 10,000 to 50,000 epochs. Results presented are based on performance of the MLP as measured by the mean squared error (MSE) of the network.

4. Results

Figure 3 below shows the MFCC and KSDE feature plots for a single participant while Figure 4 shows the feature plots for 5 combined participants. The single subject plots for both MFCC and KSDE shows clear feature separation based on different channels and frequency bands. Upon combining the extracted features for 5

participants, it appears that the normalized feature values of MFCC were quite consistent whereas the KSDE features show higher variability between subjects.

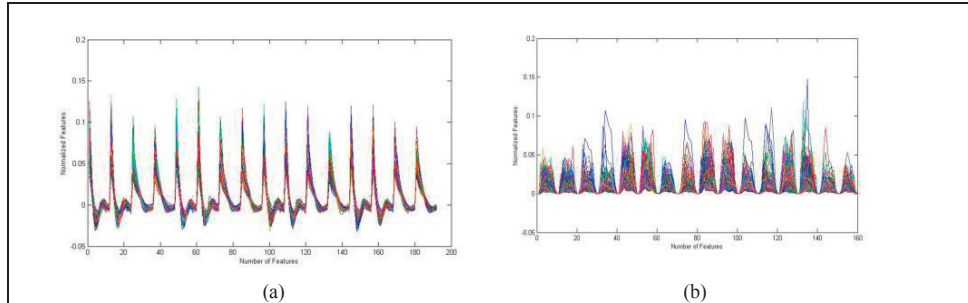


Fig.3. Single participant feature plots (Subject ID 1) for (a) MFCC, and (b) KSDE.

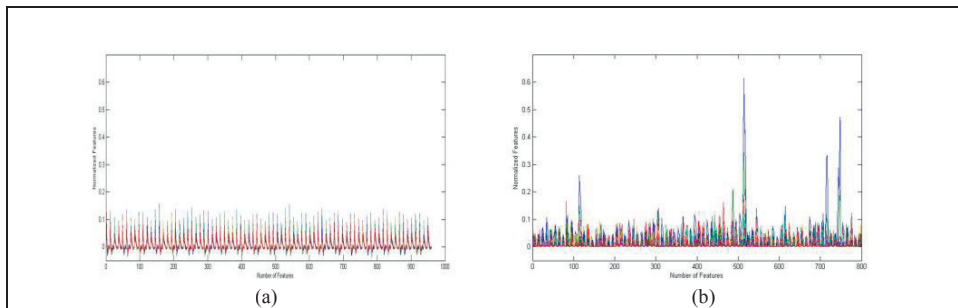


Fig. 4. Five participants feature plots for (a) MFCC, and (b) KSDE.

Apparently, results of homogenous classification in Figure 5 below shows that that the MFCC feature extraction works better with the rSASM resulting in mean squared error ranging from 0.08 to 0.15, while the performance of the MFCC-12PAC model simply hovers between 0.25 and 0.28. In contrast, the mean squared error achieved by the KSDE-rSASM model is only between 0.15 and 0.19, compared to the superior performance of the KSDE-12PAC model of 0.07 to 0.09.

In heterogeneous classification, the MFCC-rSASM and MFCC-12PAC models shows almost identical performance between 10,000 ($mse = 0.15$) and 50,000 epochs ($mse = 0.06$) when generalized against 5 participants. However, a sharp contrast can be seen between the KSDE-rSASM model and KSDE-12PAC model, with the KSDE-rSASM having mean squared error of 0.28 at 10,000 epochs compared to 0.11 at 10,000 epochs for KSDE-12PAC. However, at 50,000 epochs the performance gap between KSDE-rSASM ($mse = 0.14$) and KSDE-12PAC ($mse = 0.08$) model seems to be much closer.

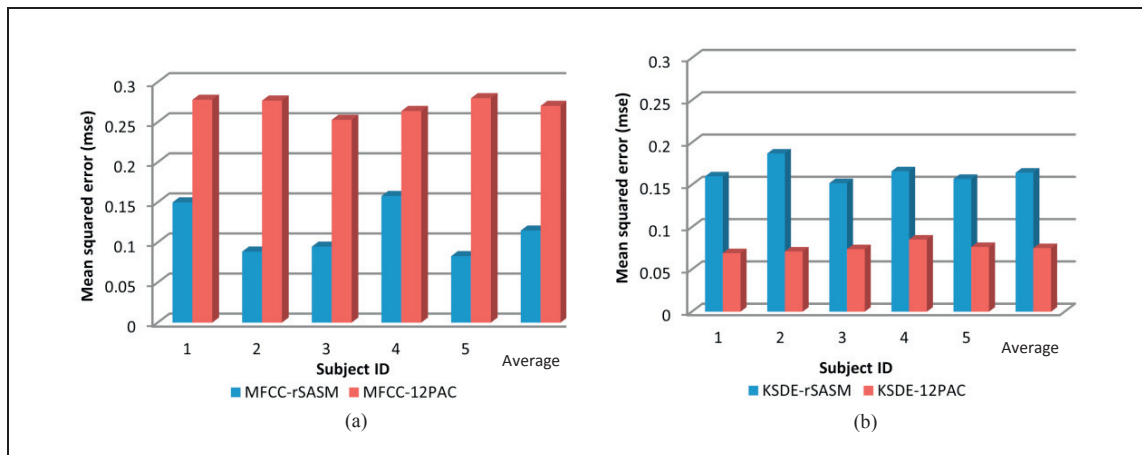


Fig. 5. Accuracy of emotion recognition system based on homogenous classification for (a) MFCC-rSASM and MFCC-12PAC models, and (b) KSDE-rSASM and KSDE-12PAC models, based on the mean squared error of the MLP network.

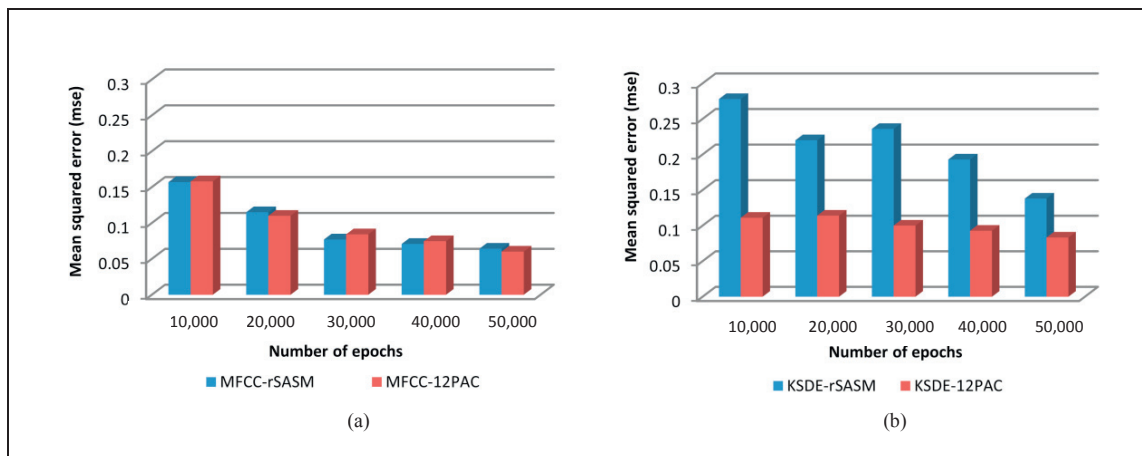


Fig. 6. Accuracy of emotion recognition system based on heterogeneous classification of 5 subjects for (a) MFCC-rSASM and MFCC-12PAC models, and (b) KSDE-rSASM and KSDE-12PAC models, based on the mean squared error of the MLP network.

5. Conclusion

In this paper, we have provided evidences that higher accuracy of the EEG emotion recognition system can be achieved with higher precision of the psychological emotion models. In essence it highlights the importance of collaborations between multidisciplinary researchers for building better solutions. The rSASM [4] is constructed based on some approximation done by researchers in computer science for interpreting the dimensional model proposed by psychologists, while the slightly detailed 12-PAC model [8] was developed by psychologists through empirical work with a large amount of data. Our analysis may be later expanded for the construction of an automated tool for the understanding of children's emotion during intervention sessions with psychologists. If such tool were to be developed, then the interpretation of emotions can be based on human physiological signals that are perceived as unbiased and much closer to biological responses, rather than relying on self-reported data that is subjective in nature and may have higher cultural contributions [21].

If the notion of gradual brain maturity is to be followed, then the homogenous classification approach using the KSDE-12PAC model (average mse = 0.08, epochs = 10,000) can be applied for achieving higher precision of emotional profiling. On the other hand, if children's acquisition of emotions are assumed to be fast and early, then heterogeneous classification with either the MFCC-12PAC (mse = 0.06, epochs = 50,000) or KSDE-12PAC (mse = 0.08, epochs = 50,000) can be employed. In such intervention sessions, real-time processing and results reporting are definitely much desirable by practitioners compared to an offline analysis approach. Higher training iterations (i.e. larger epochs) are usually at the expense of response time. It is thus our proposed system is limited to smaller training epochs in homogenous classification, while larger epochs may be set for heterogeneous classification assuming that the generalized network model is readily constructed prior to online analysis.

There are plenty of rooms for improvements of our proposed system. The dimensional models can be tested with other combinations of feature extraction and classification approaches. In addition, the performance of EEG emotion recognition system is only based on the mean squared error during the training of the MLP network. Other measures of performance may also be analyzed in the future such as the sensitivity and specificity of emotions detected by the system.

Acknowledgements

M. O. thanks Rica Frances Talon from the National Autism Society of Malaysia (NASOM) for some preliminary feedbacks on the emotional stimuli and the families of the participants of this study.

References

- [1] Murugappan M, Rizon M, et al. Time-Frequency Analysis of EEG Signals for Human Emotion Detection. In: *Proc.Kuala Lumpur 4th International Conference on Biomedical Engineering*, R. Magjarevic, Springer Berlin Heidelberg, **21**, pp. 262-65.
- [2] Ma Li, Quek Chai, Teo Kaixing, Abdul Wahab, Huseyin Abut. EEG Emotion Recognition System. In: Takeda K, et al, editors, *In-Vehicle Corpus and Signal Processing for Driver Behavior*. Springer, 2009.
- [3] Picard RW. *Affective Computing*. The MIT Press. 2000.
- [4] Kamaruddin N, Abdul Wahab. Human behavior state profile mapping based on recalibrated speech affective space model. In: *34th Annual International Conference of the IEEE EMBS*, 28 August - 1 September, 2012, San Diego, California USA.
- [5] Scherer KR. What are emotions? And how can they be measured? *Social Science Information* 2005, **44**(4), p. 693-727.
- [6] Fridja NH. The psychologists' point of view. In: Lewis M, Haviland-Jones JM, Feldman Barret L, editors. *Handbook of emotions*, 3rd Edition, p. 68-87, New York: Guilford.
- [7] Russell JA. A circumplex model of affect. *Journal of Personality and Social Psychology* 1980, **39**, p. 1161-78.
- [8] Yik M, Russell JA, Steiger JH. A 12-point circumplex structure of core affect. *Emotion* 2011, **11**(4), p. 705-31.
- [9] Ghosh-Dastidar S. Models of EEG data mining and classification in temporal lobe epilepsy: Wavelet-chaos-neural network methodology and spiking neural networks. Unpublished doctoral dissertation. The Ohio State University, Ohio, United States of America, 2007.
- [10] Gavin WJ, Davies PL. Obtaining Reliable Psychophysiological Data with Child Participants: Methodological Challenges. In: Schmidt LA, Segalowitz, SJ (editors), *Developmental Psychophysiology: Theory, Systems and Methods*, p. 424-49, New York: Cambridge University Press, 2008.
- [11] Widen SC, Russell JA. Children acquire emotion categories gradually. *Cognitive Development* 2008, **23**, p. 291-312. DOI: 10.1016/j.cogdev.2008.01.002.
- [12] Wellman, H. M., Harris, P. L., Banerjee, M., & Sinclair, A. Early understanding of emotion: Evidence from natural language. *Cognition and Emotion* 1995, **9**, p. 117-49.
- [13] Widen SC, Russell JA. A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology* 2003, **39**, 114-28. DOI: 10.1037/0012-1649.39.1.114.
- [14] Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, Knippenberg A. Presentation and validation of the Radboud Faces Database, *Cognition and Emotion* 2010, pp. 1-12.
- [15] Aron AR. Progress in Executive Functions Research: From Tasks to Functions to Regions to Networks, *Current Directions in Psychological Science* April 2008, **17**(2), p. 124-9.

- [16] Eimer M, Holmes A. Event-related brain potentials correlates of emotional face processing. *Neuropsychologia* 2007, **45**(1), p. 15-31.
- [17] Gudmunsson S, Runarsson TP, Sigurdsson S, Eiriksdottir G, Johnsen K. Reliability of quantitative EEG features. *Clinical Neurophysiology* 2007, **118**, p. 2162-71.
- [18] Krause CM, Viemerö V, Rosenqvist A, Sillanmäki L, Åström T. Relative electroencephalographic desynchronization and synchronization in humans to emotional film content: an analysis of the 4-6, 6-8, 8-10 and 10-12 Hz frequency bands. *Neuroscience Letters* 2000, **286**, pp. 9-12.
- [19] Abdul Qayoom, Othman M, Yaacob H, Abdul Wahab. EEG Affect Analysis based on KDE and MFCC, *The 2nd International Conference on Advanced Computing and Communications (ACC 2012)*, Los Angeles, CA.
- [20] Othman M, Abdul Wahab. (2010). Affective Face Processing Analysis in Autism using Electroencephalogram, *Proceeding of the International Conference on Information and Communication Technology for The Muslim World 2010*, 13-14 December 2010, Jakarta, Indonesia.
- [21] Matsumoto D, and Hyi SH. (2012). Culture and emotion: The integration of biological and cultural contributions, *Journal of Cross-Cultural Psychology* 2012, **43**(1), pp. 91-118.