Journal of Mobile Multimedia, Vol. 11, No.3&4 (2015) 296-312

EMOTION AND MOOD RECOGNITION IN RESPONSE TO VIDEO

DINI HANDAYANI, ABDUL WAHAB, HAMWIRA YAACOB

Computer Science Department, Faculty of Information and Communication Technology

International Islamic University Malaysia, Malaysia

e-mail: dini.handayani@gmail.com, abdulwahab@iium.edu.my, hyaacob@iium.edu.my

This paper presents a subject-dependent homogenous emotion recognition method using electroencephalogram (EEG) signals in response to video contents, and a correlation between emotions and moods of subjects in resting state. In the recent years, there has been a trend towards recognizing emotions invoked from watching videos. Thus, in this study, two video clips with explicit emotional contents from movies and online resources were used, and the EEG results were recorded from four subjects as they watched these clips. The best accuracies of 60.71% for valence and 63.73% for arousal were obtained using a Mel-frequency cepstral coefficients (MFCC) and multilayer perceptron (MLP). The results show that MFCC and MLP techniques are applicable in emotion recognition. The result shows that the mood can be recognized from opened eyes or closed eyes experiment of a subject. Furthermore, the results demonstrated that a positive video content can stimulate a subject into being in positive emotional state even when the subject was in bad mood. The emotional state in response to watching a video was shown to be correlated with Self-Assessment Manikin analysis.

Keywords: Affective computing, mood, emotion recognition, video clip, EEG

1 Introduction

Human emotion can be invoked from several types of stimulus, namely voice, image, video, and others [1]. In the field of affective computing, there has been a research trend towards recognizing emotions and moods elicited by watching video. Video is widely believed to be an important aspect of eliciting human emotions. When a subject watches a video, he/she would experience certain feelings and emotions based on his/her judgment of the situation which manifest through bodily and physiological cues [2]. However, a subject's emotional experience while watching a video depends on many aspects, such as personal circumstances, educational background, race, sex, and age [3].

Several measurement techniques were used to recognize emotions and moods. Specifically, by using self-assessment form, electroencephalogram (EEG), pupillary response and gaze distance, facial expression, electromyography (EMG), electrocardiogram (ECG), Galvanic Skin Response (GSR), and respiration pattern. The measurement tools (for emotion and mood recognition) used for this study were EEG and Self-Assessment Manikin (SAM). In brief, an EEG can denote a subjective emotional state based on the subject's own experiences, which is reflected from the brain's electrical activities [4]. It has become a prominent technique to capture brain electrical activities as it is less costly compared to others, such as Functional Magnetic Resonance Imaging (FMRI) and Positron Emission Tomography (PET). Furthermore, SAM is a commonly used method to measure emotional state by using a self-reporting form.

Another essential point is the data analysis methods that were employed to develop computational model on emotions and moods. Some works had been done using statistical and machine-learning techniques. Furthermore, for feature extraction techniques, studies had been done using Power Spectral Density (PSD), Common Spatial Patterns (CSP), Hybrid Adaptive Filtering-Higher Order Crossing (HAF-HOC), Affective Audio Visual Words (AAVWs), Short Time Fourier Transform (STFT), and so on. For classification techniques, some works had been done using Support Vector Machines (SVM), 2D Gaussian probability, Bayesian Linear Discriminant Analysis (BLDA), Support Vector Regression (SVR), Adaptive Network-based Fuzzy Inference System (ANFIS), Naïve Bayes (NB), Linear Discriminant Analysis (LDA), Hidden Markov Models (HMM), Hierarchical HMM (HHMM), Latent Topic Driving Model (LTDM), Wavelet, K Nearest Neighbor (KNN), Random Forest (RF), Quadratic Discriminant Analysis (QDA), Multilayer Perceptron (MLP), Radial Basis Function Neural network (RBFN), and so on. In this study, Mel-frequency Cepstral Coefficients (MFCC) and MLP were utilized for feature extraction and classification.

The aim of this study is to recognize emotions invoked from watching a video and their correlation with the moods of subjects in resting state. Emotion and mood can mutually influence each other. Emotion, if strong and deep enough, will turn into mood. It is therefore important to understand the mood of a subject and its relation to his/her emotion.

2 Related Work

In this section we briefly explore the limitations of current research work and contrast this with commercially accepted development practices. We then discuss anecdotal evidence that supports an iterative development process incorporating client-developer joint exploration of partial designs to facilitate the development of client understanding of their needs.

Several researchers have studied the influence of videos on human emotion, as shows in Table 1. Audio-visual information from videos plays an important role in human emotion. Therefore, some of the previous studies focused on emotional tagging of the video [5]–[8]. Emotional tagging plays an important role for video content profiling.

Recently, empirical research on emotion and mood recognition is characterized by a wide variety of methodologies. Emotion and mood recognition can be done through explicit and implicit approaches. Typical explicit measurement is by using self-assessment forms, while the implicit approach is an automatic inference of emotional information based on the measurement of the behaviour and physiological signals from the human [9].

Applying explicit measurement, Irie et al., 2010, presented a novel method for classification of movie scenes based on emotional content. The experimental study used self-assessment of 16 adult participants to analyze 206 movie scenes [10]. Sanchez-Cortes et al., 2013, presented an emotion recognition method from conversational videos by using self-assessment responses of participants. The finding of the study revealed that to generate a mood model, it is important to know the context of the video [11]. Baveye et al., 2015, utilized SAM to understand the human emotional response to video content. In this study, 1,517 adults participated to analyze 9,800 video clips. The result was a video database for affective content analysis, namely LIRIS-ACCEDE, with valence and arousal as affect representation [12].

Other than that, there were also other researches done on emotion and mood recognition through implicit approaches using different physiological conditions, such as EEG [13]–[15], pupillary response and gaze distance [7], [8], [16], facial expression [4], [8], EMG [17], ECG [2], GSR [2], and respiration pattern [2].

Park et al., 2011, developed a real-time emotion monitoring system based on EEG recording data from 34 subjects while watching emotional video [13]. Likewise, [14] presented a human emotion recognition method using EEG signals as subjects watched video clips. The classification used five-fold cross validation. Additionally, the experimental study also suggested the possibility of differentiating human emotions by linear and non-linear features. [15], in their study, used EEG signals for emotion recognition. The experimental study was done on five subjects. The result showed that emotional states can be recognized by frontal and parietal EEG signals. Soleymani et al., 2012, presented a subject-independent emotion recognition method in response to watching 20 videos. The experimental study was done using EEG, pupillary response, and gaze distance from 24 participants. The best accuracy was achieved using modality fusion and SVM [7]. The previous study that utilized EEG and other multi-modal approach was done in [4]. Feature- and decision-level fusions were applied for feature extraction techniques, and the classification was done using regression.

Utilizing explicit and implicit approaches concurrently, Bastos-Filho et al., 2012, conducted a study using EEG signals and SAM forms. The experiment was done using Database for Emotional Analysis using Physiological Signals (DEAP). SAM forms were used concurrently with four-channel EEG signals [18]. Koelstra et al., 2012, presented a multimodal data set for analyze human emotion response to 40 music videos. The affective representation used arousal, valence, and dominance [19]. Similarly, the relationship between EEG signals and human emotions was studied in [20]. The emotion classification was based on positive and negative emotions. The participants were told to fill in the SAM forms after watching video clips. Additionally, previous study done in [21] presented the correlation between self-assessments and EEG signals. The experimental study was conducted to recognize human emotion while watching 20 music video clips. The experimental study was recorded from six subjects to perceive the levels of valence and arousal. The result of the study showed that the use of music video can elicit human emotion from video and audio modalities. Finally, Lee et al., 2014, proposed an emotion recognition system for understanding the human emotions while watching a video clip using EEG and questionnaire. The experimental study applied Independent Component Analysis (ICA) to eliminate artifact. STFT and ANFIS were utilized as feature extraction and classification techniques [3].

Several studies were done to recognize emotion and mood using other methods. Katsimerous et al., 2014, proposed a model for mood recognition from a sequence of emotions while watching a video retrieved from well-established database namely HUMAINE. ANVIL video annotator was utilized to rate the video emotion [22]. [17], used EMG and self-assessed form to recognize three emotional videos from 34 participants. The result of the study showed that emotional states can be classified with high accuracy using cross-validation method. [2] proposed ECG, GSR, and skin temperature data for recognizing emotion while watching emotional video. The affective representatives for this research were valence and arousal.

Table 1: Taxonomical table on emotion and mood recognition in response to video stimuli

No	Source	Stimuli	Number of Subject	Affect Representation	Data Analysis & Accuracy
1	[21]	Music video	6 subjects	V, A	PSD, CSP, SVM V = 58.8%, A = 55.7%
2	[2]	64 short video clips	7 adults	V, A	2D Gaussian probability
3	[6]	24 video clips	8 (male) adults	Discreet	Bayesian linear discriminant analysis 75%
4	[12]	9,800 films clips	1517 adults	V, A	SVR
5	[3]	Movie clip	12 adults	V	ANFIS = 76.25% NB = 59.95% SVM = 78.45%
6	[10]	206 movie scenes	16 adults	Discreet	LDA-AAVW: 77.5%, SVM-AAVW: 78.3%, HMM-AAVW: 80.7%, HHMM-AAVW: 81.3%, LTDM-AAVW: 83.4%, HMM-AAVW: 81.9%, LTDM-AAVW: 85.5%
7	[7]	Movie video	24 adults	V, A	PSD, SVM V = 68.5%, A = 76.4%
8	[4]	Video clips	24 subject	V, A	PSD, HMM V = 80%, A = 80%, Control = 86.7%
9	[22]	HUMAINE video database	6 adults (expert)	V, A	Statistic = 62%
10	[13]	Video	34 adults	Discreet	Statistic
11	[14]	Video	20 adults	Discreet	Wavelet-KNN 83.04% Wavelet-LDA 79.17%
12	[15]	Video	5 adults	Discreet	SVM 66.5%
13	[16]	Video SEMAIN database	36 adults	V, A	-
14	[11]	Video conversation	5 adults	Discreet	SVM & RF = 68.5%
15	[18]	Video	32 adults	V, A	PSD-KNN:70.1%, HOC-KNN:69.59%
16	[19]	Video	12 adults	V	ANFIS 76.25% NB 59.94 SVM 78.45%
17	[8]	20 videos	27 adults	V, A	HMM
18	[17]	6 videos	37 adults	Discreet	LDA, QDA, MLP, RBFN and KNN 84.5%
19	[20]	Movie video	6 adults	V, A and D	LSD, PCA, SVM = 87.53%

A. Yazdani et al., 2009 [6] proposed a brain computer interface (BCI) based on P300 to generate emotional tagging based on six basic emotions proposed by Ekman. The experimental study was done on 12 male adult participants, with eight of them for training, and the other four participated in testing. The experimental study achieved 75% accuracy on selecting tags. In [8], face-recording videos, audio signals, eye gaze, peripheral/central nervous system physiological signals, and self-assessments were studied. The affective representation is based on valence, arousal, and dominance scale. Moreover, [16] examined how emotion feedback influences the emotion awareness and gaze behaviour. The experimental study was done on 36 participants that wore an eye-tracker each while watching 12 selected videos from SEMAINE database. The result from the experimental study showed that arousal scores provided better insights than that of valence. As the aforementioned studies did not use the EEG and SAM forms, literature review is still needed to determine the existence of other methods that may be relevant in future research.

3 Material And Methods

3.1 Emotions and Mood Model

The most straightforward way to represent an emotion is by using categorical approach or discrete labels, such as ‘anger’, ‘contempt’, ‘disgust’, ‘fear’, ‘sad’, ‘surprise’, and ‘happy’. Psychologists frequently represent emotions and moods in an n-dimensional space. Russell [23] proposed a two-dimensional affective space model for measuring emotions known as circumplex model of affect. It is composed of valence and arousal. Valence ranges from negative (unpleasant) to positive (pleasant) while arousal ranges from calm (low) to excited (high). As shown in Figure 1, the corresponding dimensions (valence and arousal) are illustrated on a Cartesian coordinate space.

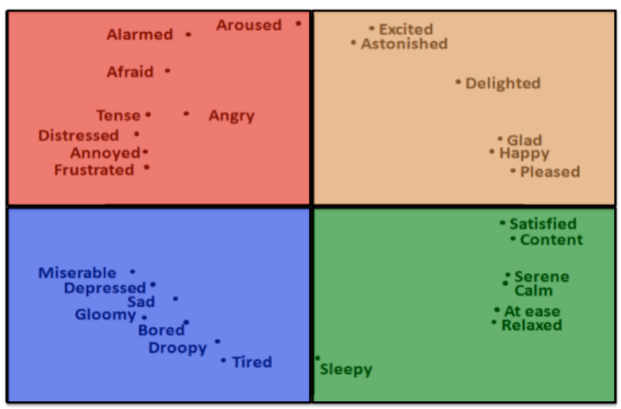


Figure 1: Circumplex Model of Affect from [23] *with the emotional state color* [24]*, whereby valence is scaled on the x-axis and arousal on y-axis*

Bialoskorski et al. (2009) labeled emotional states with colours [24], as illustrated in Figure 1. Happy emotional state, indicated in orange, is defined as having positive level of valence and high level of arousal. Calm emotional state, indicated in green, is defined as having positive level of valence but low level of arousal. Sad emotional state or depression, indicated in blue, is defined as having negative level of valence and low level of arousal. Fear or anxiety, indicated in red, is defined as having negative level of valence but high level of arousal.

3.2 Experiment Protocol and Data Collection

Four young and healthy undergraduates (two males and two females) with different educational and cultural backgrounds volunteered to participate in this experiment at International Islamic University of Malaysia (IIUM). Their ages varied between 22 to 33 years old, with mean (M) of 26.5 years old and standard deviation (SD) of 5.44 years. The experiments were carried out in a laboratory environment with controlled temperature and illumination. This research was formally approved by the Ethics Committee IIUM.

The subjects were briefed about the experiment and their rights through a consent form and with a verbal explanation. Subjects were explained about the interface before the experiment and during setup time. As a reference, the subjects were shown four emotions, namely happiness, calmness, sadness, and fear by using the International Affective Picture (IAPS) concurrently with Bernard Bouchard's musical clips from [25]; a minute for each emotion. The subjects were asked to avoid moving their bodies when they watch a 90-second video clip. The video clip was played from a data set (pcbdgdb: available online from http://PCBDG.IIUM.EDU.MY).

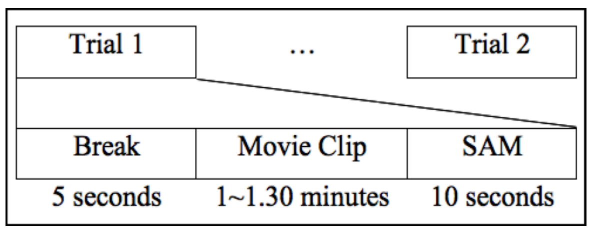


Figure 2: Each trial was conducted with a five-second break and then continued by a video. Subjects filled in the SAM forms at the end of each trial. There were two trials in each experimental session.

The time interval between the start of a trial and the end of the self-reporting phase was approximately two minutes. The whole protocol took 10 minutes on average, in addition to 5 minutes of setup time (see Figure 2). The experimental visual stimuli employed in the present study consist of video clips selected from movies and online resources.

4 The Eeg Emotion And Mood Recognition System

4.1 EEG Signals

EEG signals were recorded using EPOC headset manufactured by Emotiv Inc [26]. The EEG headset, as shown in Figure 2, is used to read a person's brain waves from the surface of the scalp. The Research Edition SDK by Emotiv Systems is a research headset; a 14-channel (plus CMS/DRL references, P3/P4 locations) high-resolution wireless neuroheadset with neuro-signal acquisition and processing functions.

Channel names based on the International 10-20 locations are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. The Research Edition SDK also includes a proprietary software toolkit that contains the APIs and detection libraries. For the purpose of this study, an 8-channel EEG with a sampling rate of 128 Hz was used, as shown in Figure 3 in red.

Table 2: Types of Brainwave Signals

Bands	Frequency	Descriptions
Delta	0.5 Hz – 4 Hz	Deep sleep [31].
Theta	4 Hz – 8 Hz	Drowsiness and fatigue due to monotonous task [32], control of working memory process [33].
Alpha	8 Hz – 13 Hz	Cognitive control [34], creative thinking [35].
Beta	13 Hz – 30 Hz	Alertness [36], phonological tasks [37].

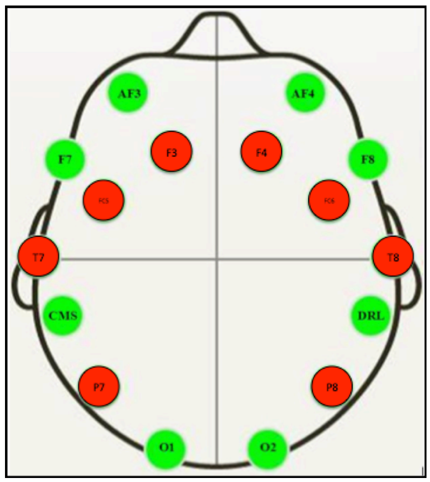


Figure 3: The Emotiv layout for 14 EEG channels in addition to two reference electrodes. Retrieved from Emotiv website (https://emotiv.com/).

4.2 Preprocessing and Feature Extraction

In the preprocessing stage, EEG signals were filtered to eliminate noises and artifacts. Brain waves, which are represented by EEG signals, are commonly ranged into four bands: delta, theta, alpha, and beta. Each of the frequency bands is observed as products of different brain tasks, with some depicted in Table 2.

To extract features from the filtered signals, MFCC was adopted. Twelve linearly spaced mel filter banks were computed for each frequency band and eight EEG channels were used to obtain the mel cepstrum. As a result, 192 features were extracted for each instance.

4.3 Emotion Classification

MLP was employed to construct a classifying network that output categories of emotion based on valence and arousal value inputs. For the purpose of classification, the following network parameters were established in Table 3.

Table 3: Parameters for MLP

Parameters	Values
No. of hidden layer	2
No. of neurons in hidden layer	20, 20
No. of neurons in output layer	1
Mean-square error goal	0.1

5 Experimental Results

The accuracy of the experiment are shown in Figure 4. It is similar to previous studies done in [2], [12], [21], using valence and arousal as affective representation. By separating valence and arousal scores, it yields better accuracy compared to the combination of valence and arousal scores. For the valence approach, M is 54.23% and SD 5.61%. For the arousal dimension, M is 53.87% and SD 8.77%.

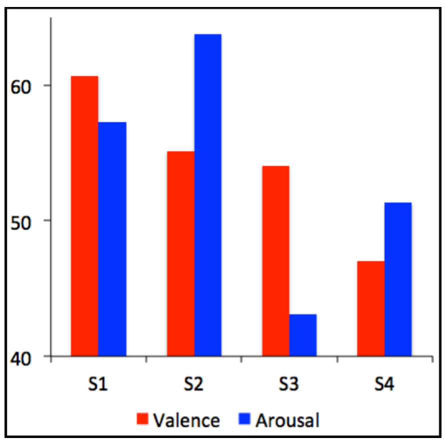


Figure 4: Accuracy of Subjects' Identification

The performance of classification was validated using the five-fold cross validation in the same way done in [14], [17]. Figure 4 presents the results of five-fold cross validation using features that were extracted using MFCC and MLP as a classification technique.

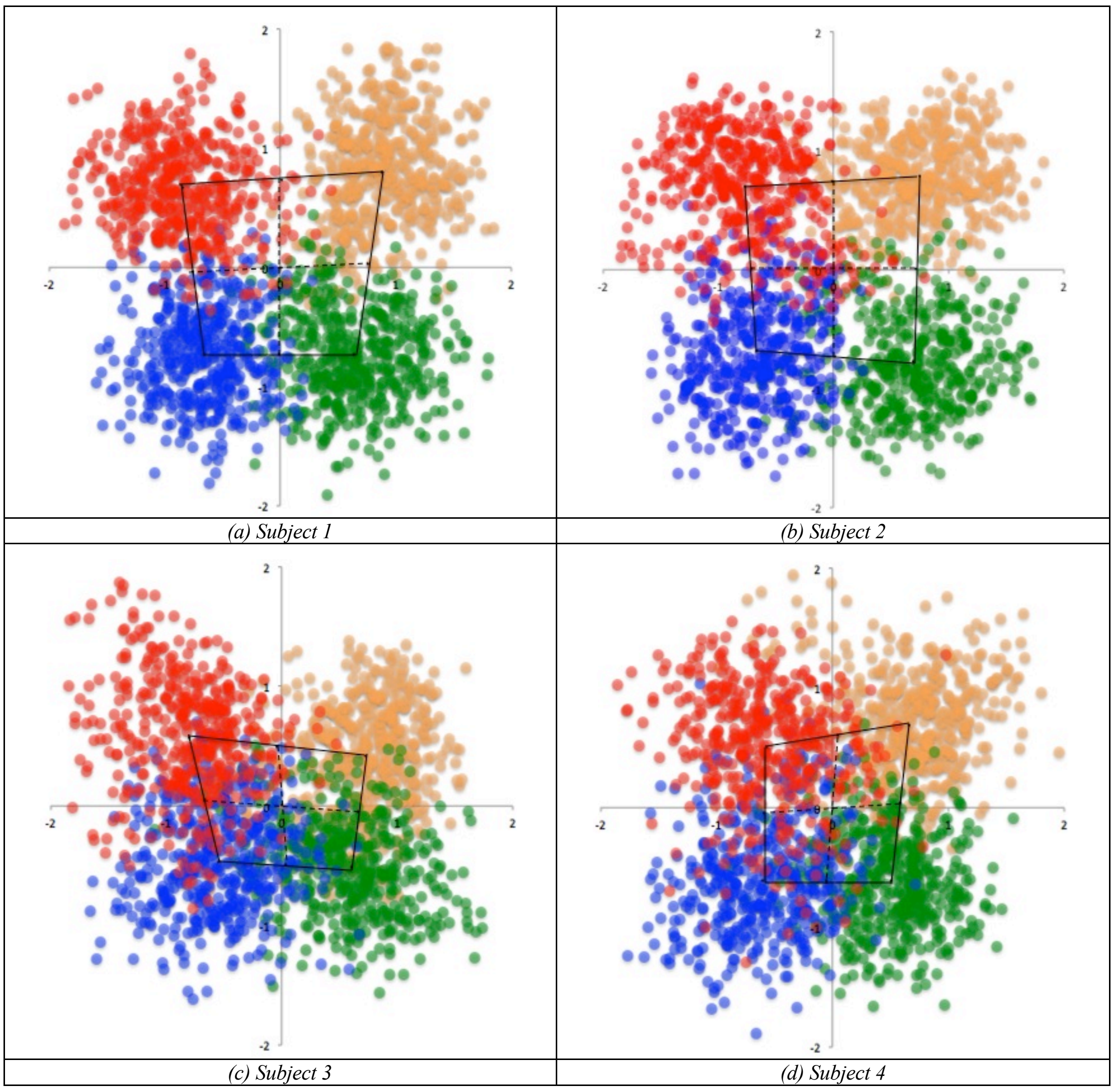
5.1 Emotion Recognition and EEG

The emotional states are represented on the Cartesian coordinates with x-axis representing emotional valence and y-axis representing emotional arousal. Table 4 presents the emotions' intended output values as outlined by Kamaruddin and Abdul Wahab 2012 [27].

Different subjects have different level of emotion. Here, the recalibrated affective space model is proposed to get the new center point for each subject. The recalibration technique was previously proposed by Kamaruddin and Abdul Wahab (2012) to recognize culturally-influenced speech emotion, namely recalibrated Speech Affective Space Model (rSASM) [27]. Figure 5 presents the emotion distribution for each subject.

Table 4: Intended output for different emotions (source: [27]*)*

Emotion	Valence	Arousal
Happy	+1	+1
Calm	+1	−1
Sad	−1	−1
Fear	−1	+1
Neutral (center)	0	0



(a) Subject 1 *(b) Subject 2*

(c) Subject 3 *(d) Subject 4*

Figure 5: Emotions distribution with recalibration the center point.

As a result of emotion distribution, every emotion is quite well distributed in each emotion quadrant as shown in Figure 5. Every subject generated a new set of emotion quadrants. However, the center point of each recalibrated emotion distribution is near to the original center point, which is (0, 0). Therefore, for all subjects in this study, there was no change in the center point. Every emotion visualization for all the subjects would refer to the original affective space. The center point is 0, the valence positive is +1, valence negative is −1, as well as high arousal is +1 and low arousal is −1.

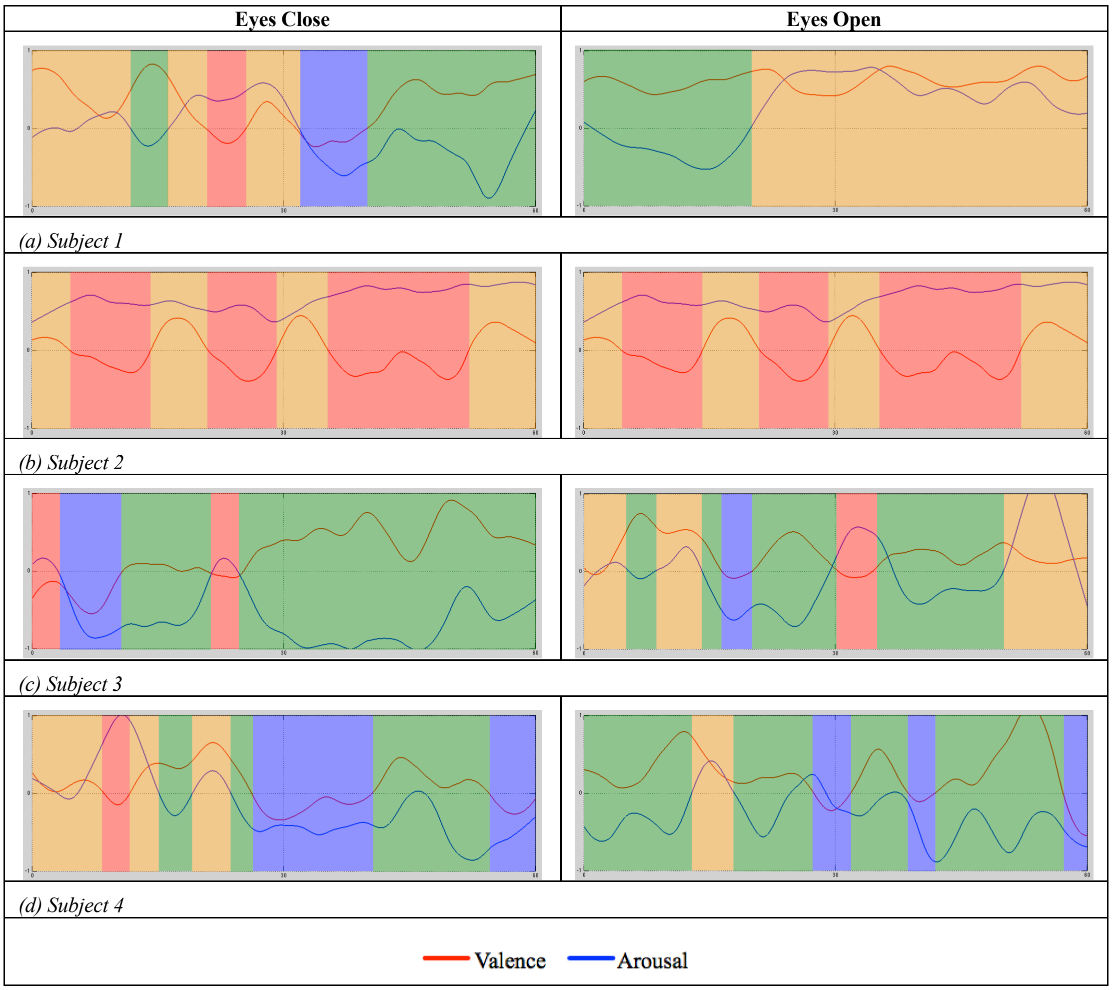


Figure 6: Mood Visualization

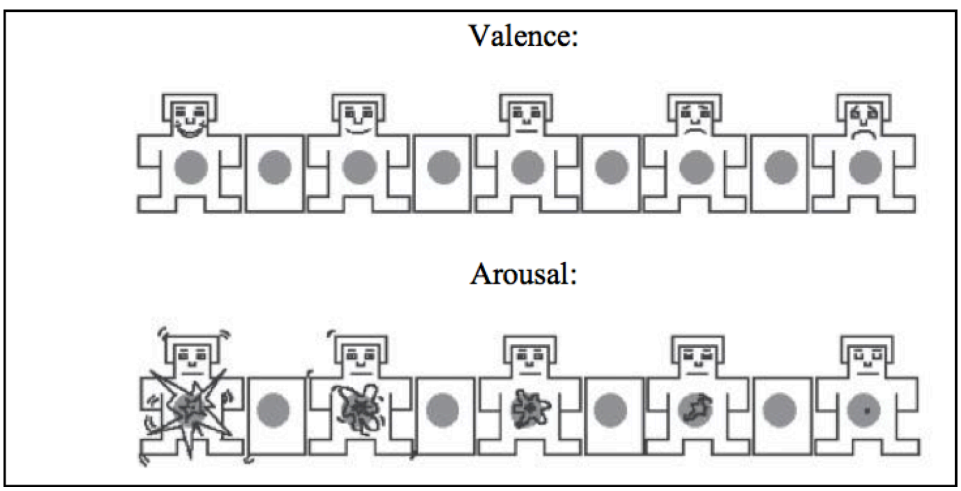
Figure 6 demonstrates the mood visualization of four subjects under resting state; when they had their eyes closed and when they had their eyes open. To get the mood under resting state, moving average was calculated for both valence and arousal as formulated in [28]. With the smoothing average of five seconds, the result is presented in Figure 6.

Table 5: Mood Recognition under Resting State, Happy (H), Calm (C), Sad (S), Fear (F)

Subject	**Gender**	**Resting State**	**Start**	**End**	**Dominant**	**Initial Mood**
1	F	EC	H	C	C	Positive
		EO	C	H	H	
2	M	EC	H	H	F	Negative
		EO	H	F	F	
3	F	EC	F	C	C	Positive
		EO	H	H	C	
4	M	EC	H	S	C	Positive
		EO	C	S	C	

Mood is categorized into three types, specifically positive, negative, and neutral mood as described in [29]. The categorization is based on valence scale. Positive valence that shows positive mood is defined as ‘good mood’, while negative valence that shows negative mood is defined as ‘bad mood’.

Neutral valence is defined as ‘neutral mood’. Resting states of the subjects are believed to reflect their initial mood during the experiment.



SCORING: 9 8 7 6 5 4 3 2 1

Figure 7: SAM

It is very important to know the initial mood before the start of the experiment. It would show the mood changes towards the specific emotion that are elicited during the induction task. The initial mood was inferred from the dominant emotions that were derived from when the subjects had their eyes closed (EC) and eyes opened (EO). Mood visualization from Figure 6 are summarized in Table 5. As shown in

Table 5, Subject 1, 3, and 4 had the positive valence as dominant emotion. Therefore, for these subjects, the initial moods were positive. In other words, they were in good mood.

On the other hand, Subject 2 had dominant negative emotion. Thus, the initial mood was negative; Subject 2 was in a bad mood.

5.2 Video Emotion Recognition and Self-Reports

The main interest of this research is the correlation between the subjects' initial moods and their change of moods in response to video content. For the purpose of this research, two positive videos were shown to each subject, namely happy and calm videos. The happy video used for this research was "Baby Laughing Hysterically at Ripping Paper". The calm video was "Havasupai Indian Waterfall Relaxation". The present study utilized both EEG signals and SAM forms similar to previous studies done in [18], [20], [21]. In every trial, EEG signals were recorded as the subjects watched the video clip. After watching it, the subjects filled in the SAM forms.

SAM is a commonly used technique to measure emotional states [30]. It is a self-reporting affective state measurement, using cartoon-like manikin (see Figure 7) to plot basic emotions on the affective space. A nine-point pictorial scale was utilized for the purpose of this study. In Figure 7, similar to previous study done in [12], two sets of manikin were used. The first set was used for valence scoring, with the range from nine (happy) to one (sad). The second set was for arousal scoring, with the range from nine (active) to one (passive). A SAM is defined as '*Happy*' when the levels of valence and arousal are both 5 and above:

$$(valence \geq 5) \cap (arousal \geq 5) \tag{1}$$

'*Calm*' is when the levels of valence is 5 and above and arousal below 5:

$$(valence \geq 5) \cap (arousal < 5) \tag{2}$$

'*Sad*' is when the levels of both valence and arousal are below 5:

$$(valence < 5) \cap (arousal < 5) \tag{3}$$

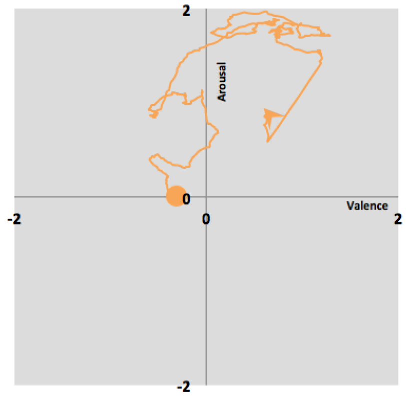
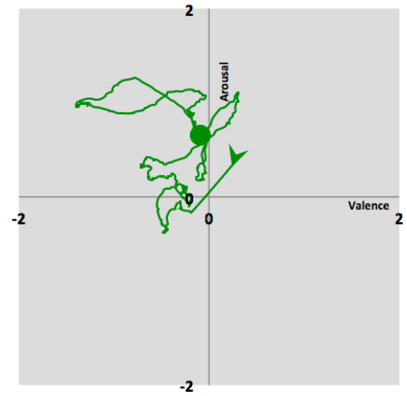
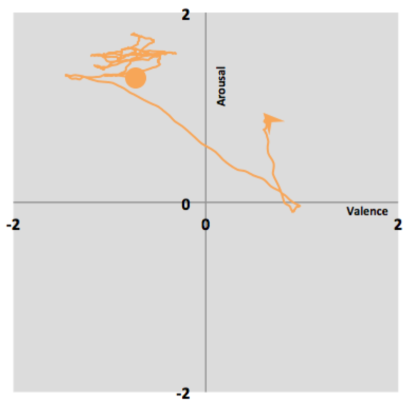
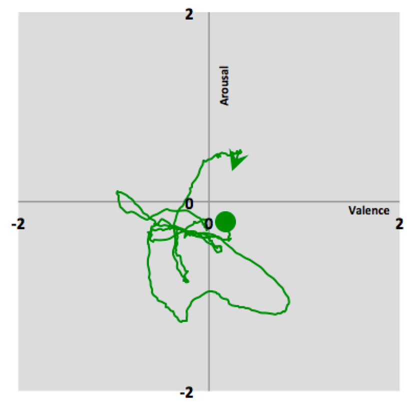
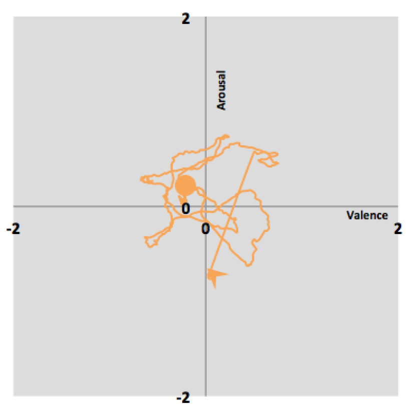
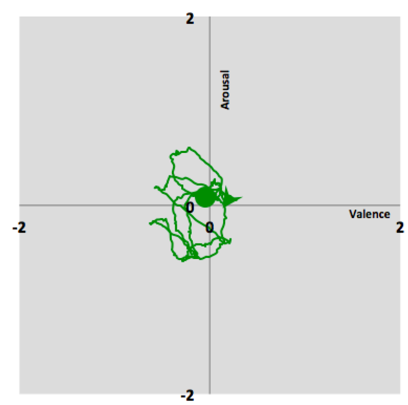
'*Fear*' is when the levels of valence is below 5 and arousal 5 and above:

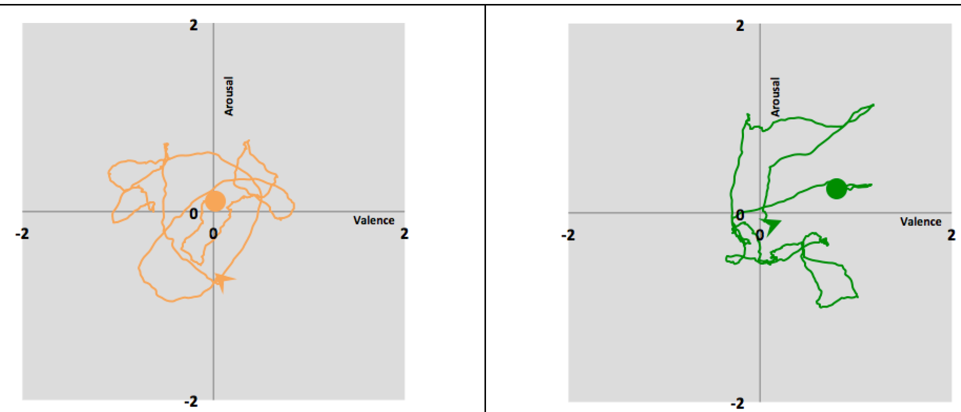
$$(valence < 5) \cap (arousal \geq 5) \tag{4}$$

Figure 8 presents the results of the subjects' mood recognition as they watched the video clips. The happy emotion video is indicated in orange, while calm emotion video is indicated in green.

Table 6: Mood Recognition while watched the video clips

Subject	**Initial Mood**	**Video Emotions**	**Start**	**End**	**SAM**
1	Positive	VH	F	H	H
		VC	F	H	H
2	Negative	VH	F	H	C
		VC	C	H	C
3	Positive	VH	F	C	H
		VC	F	H	H
4	Positive	VH	H	C	H
		VC	H	C	H

<table>
<tr><th>VH</th><th>VC</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>(a) Subject 1</td><td></td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>(b) Subject 2</td><td></td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>(c) Subject 3</td><td></td></tr>
</table>



(d) Subject 4

Happy Calm

Figure 8: Mood Recognition Visualization in Response to Video

The EEG results from the subjects as they watched the video clips were compared to their SAM results, as shown in Table 6. As expected in the results, all the subjects' EEG could recognize the videos as having positive emotional contents. As for the SAM results, it shows that all subjects could recognize the happy and calm videos as having positive emotional contents as well.

It is observed that all the results were in positive valence. However, the EEG and SAM results did not agree with each other completely in terms of 'happy' or 'calm' due to different arousal levels between the pairs of EEG and SAM results. It is also observed that no matter what the initial moods of the subjects were (whether they were positive or negative), the positive video clips changed their emotions to positive valence by the end of the videos.

The results also demonstrated that there was no significant difference between subjects with different cultural background, gender, and age [3].

6 Conclusions

This paper has shown a good performance of a subject-dependent homogenous emotion recognition method using subjects' EEG signals when watching video. Although the results were based on a small video data set, they show that the video clips can change the subjects' moods into the emotions conveyed by the videos. Additionally, the results from the study show that there is no significant difference between the subjects' EEG signals and their self-assessment results.

Even though the classification results were based on a small number of subjects due to experimental limitations, the promising accuracies suggest that the experiment can be scaled to a large number of samples. Experimental results have shown that MFCC and MLP produce high accuracy and consistency. They can be tested with other feature extraction techniques and more sophisticated classifiers in the future work to get more robust results.

This research has open issues that need to be considered in the future; a larger set of videos with various emotional contents can be considered to increase the generalization of the emotion recognition in response to video contents.

Acknowledgment

This work is supported by Fundamental Research Grant Scheme (FRGS) funded by the Ministry of Higher Education (Grant code: FRGS14-137-0378).

References

1. R. Khosrowabadi, A. Wahab, K. K. Ang, and M. H. Baniasad, "Affective Computation on EEG Correlates of Emotion from Musical and Vocal Stimuli," *Proceeding Int. Jt. Conf. Neural Networks, Atlanta, Georg. USA*, pp. 1590–1594, 2009.

2. J. J. M. Kierkels, M. Soleymani, and Thierry Pun, "Queries and Tags in Affect-Based Multimedia Retrieval," *Int'l. Conf.Multimedia Expo, Spec. Sess. Implicit Tagging (ICME 2009)*, pp. 1436–1439, 2009.

3. G. Lee, M. Kwon, S. Kavuri Sri, and M. Lee, "Emotion Recognition Based on 3D Fuzzy Visual and EEG Features in Movie Clips," *Neurocomputing*, vol. 144, pp. 560–568, Nov. 2014.

4. S. Koelstra and I. Patras, "Fusion of facial expressions and EEG for implicit affective tagging," *Image Vis. Comput.*, vol. 31, no. 2, pp. 164–174, Feb. 2013.

5. J. J. M. Kierkels, B. B. A, R. De Drize, and C. H. Carouge, "A Bayesian Framework for Video Affective Representation Mohammad Soleymani Guillaume Chanel Thierry Pun Computer Vision and Multimedia Laboratory , Computer Science Department," 2009.

6. A. Yazdani, J.-S. Lee, and T. Ebrahimi, "Implicit Emotional Tagging of Multimedia Using EEG Signals and Brain Computer Interface," *Proc. first SIGMM Work. Soc. media - WSM '09*, p. 81, 2009.

7. M. Soleymani, M. Pantic, and T. Pun, "Multimodal Emotion Recognition in Response to Videos," *IEEE Trans. Affect. Comput.*, vol. 3, no. 2, pp. 211–223, Apr. 2012.

8. M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A Multimodal Database for Affect Recognition and Implicit Tagging," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 42–55, Jan. 2012.

9. J. Broekens and W.-P. Brinkman, "AffectButton: A method for Reliable and Valid Affective Self-report," *Int. J. Hum. Comput. Stud.*, vol. 71, no. 6, pp. 641–667, Jun. 2013.

10. G. Irie, T. Satou, A. Kojima, T. Yamasaki, and K. Aizawa, "Affective Audio-Visual Words and Latent Topic Driving Model for Realizing Movie Affective Scene Classification," *IEEE Trans. Multimed.*, vol. 12, no. 6, pp. 523–535, Oct. 2010.

11. D. Sanchez-Cortes, J.-I. Biel, S. Kumano, J. Yamato, K. Otsuka, and D. Gatica-Perez, "Inferring Mood in Ubiquitous Conversational Video," *Proc. 12th Int. Conf. Mob. Ubiquitous Multimed. - MUM '13*, pp. 1–9, 2013.

12. Y. Baveye, E. Dellandréa, C. Chamaret, and L. Chen, "LIRIS-ACCEDE: A Video Database for Affective Content Analysis," *Affect. Comput. IEEE Trans.*, pp. 1–14, 2015.

13. K. Park, H. Choi, and K. Lee, "Emotion Recognition Based on The Asymmetric Left and Right Activation," *Int. J. Med. Med. Sci.*, vol. 3, no. June, pp. 201–209, 2011.

14. M. Murugappan, R. Nagarajan, and Sazali Yaacob, "Combining Spatial Filtering and Wavelet Transform for Classifying Human Emotions Using EEG Signals," *J. Med. Biol. Eng.*, vol. 31, no. 1, pp. 45–51, 2011.

15. X. Wang, D. Nie, and B. Lu, "EEG-Based Emotion Recognition Using Frequency Domain Features and Support Vector Machines," *Neural Inf. Process.*, pp. 734–743, 2011.

16. F. Ringeval, A. Sonderegger, B. Noris, A. Billard, J. Sauer, and D. Lalanne, "On the Influence of Emotional Feedback on Emotion Awareness and Gaze Behavior," *2013 Hum. Assoc. Conf. Affect. Comput. Intell. Interact.*, pp. 448–453, Sep. 2013.

17. V. Kolodyazhniy, S. D. Kreibig, J. J. Gross, W. T. Roth, and F. H. Wilhelm, "An affective computing approach to physiological emotion specificity: toward subject-independent and stimulus-independent classification of film-induced emotions.," *Psychophysiology*, vol. 48, no. 7, pp. 908–22, Jul. 2011.

18. T. F. Bastos-Filho, A. Ferreira, A. C. Atencio, S. Arjunan, and D. Kumar, "Evaluation of Feature Extraction Techniques in Emotional State Recognition," *2012 4th Int. Conf. Intell. Hum. Comput. Interact.*, pp. 1–6, Dec. 2012.

19. S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. (Yiannis) Patras, "DEAP : A Database for Emotion Analysis Using Physiological Signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, 2012.

20. D. Nie, X. Wang, L. Shi, and B. Lu, "EEG-based Emotion Recognition during Watching Movies," *Proc. 5th Int. IEEE EMBS Conf. Neural Eng.*, pp. 667–670, 2011.

21. S. Koelstra, A. Yazdani, M. Soleymani, C. Mühl, J.-S. Lee, A. Nijholt, T. Pun, T. Ebrahimi, and I. Patras, "Single trial classification of EEG and peripheral physiological signals for recognition of emotions induced by music videos," *Brain informatics*, pp. 89–100, 2010.

22. C. Katsimerou, J. Redi, and I. Heynderickx, "A Computational Model for Mood Recognition," *22nd Int. Conf. UMAP 2014, Aalborg, Denmark*, vol. 8538, pp. 122–133, 2014.

23. J. Russell, "A circumplex model of affect.," *J. Pers. Soc. Psychol.*, 1980.

24. L. S. S. Bialoskorski, J. H. D. . Westerink, and E. L. van den Broek, "Mood Swings: An affective Interactive Art System," *ICST Inst. Comput. Sci. Soc. Informatics Telecommun. Eng. 2009*, pp. 181–186, 2009.

25. R. Khosrowabadi, H. C. Quek, A. Wahab, and K. K. Ang, "EEG-based Emotion Recognition Using Self-Organizing Map for Boundary Detection," *2010 20th Int. Conf. Pattern Recognit.*, pp. 4242–4245, Aug. 2010.

26. Emotiv, “Emotiv Epoc,” 2014. [Online]. Available: http://www.emotiv.com/epoc.php. [Accessed: 17-Jun-2015].

27. N. Kamaruddin and A. Wahab, “Human behavior state profile mapping based on recalibrated speech affective space model.,” *Conf. Proc. 34th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc.*, pp. 2021–4, Jan. 2012.

28. D. Handayani, H. Yaacob, A. W. A. Rahman, W. Sediono, and A. Shah, “Computational Modeling of Mood from Sequence of Emotions,” *3rd Int. Conf. Internet Serv. Technol. Inf. Eng. 2015*, pp. 154–158, Feb. 2015.

29. Y. Guoliang and W. Zhiliang, “Affective computing model based on emotional psychology,” *Affect. Comput. Model Based Emot. Psychol.*, pp. 251–260, 2006.

30. M. Bradley and P. J. Lang, “Measuring Emotion: The Self-Assessment Manikin and The Semantic Differential,” *J. Behav. Ther. Exp. Psychiat.*, vol. 25, no. I, 1994.

31. K. Šušmáková, “Human Sleep and Sleep EEG Institute of Measurement Science , Slovak Academy of Sciences,” vol. 4, pp. 4–5, 2004.

32. B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, “Using EEG spectral components to assess algorithms for detecting fatigue,” *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2352–2359, Mar. 2009.

33. P. Sauseng, B. Griesmayr, R. Freunberger, and W. Klimesch, “Control mechanisms in working memory: a possible function of EEG theta oscillations.,” *Neurosci. Biobehav. Rev.*, vol. 34, no. 7, pp. 1015–22, Jun. 2010.

34. B. Zoefel, R. J. Huster, and C. S. Herrmann, “Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance.,” *Neuroimage*, vol. 54, no. 2, pp. 1427–31, Jan. 2011.

35. A. Fink, B. Graif, and A. C. Neubauer, “Brain Correlates Underlying Creative Thinking: EEG Alpha Activity in Professional vs. Novice Dancers.,” *Neuroimage*, vol. 46, no. 3, pp. 854–62, Jul. 2009.

36. J. Kamiński, A. Brzezicka, M. Gola, and A. Wróbel, “B Band Oscillations Engagement in Human Alertness Process.,” *Int. J. Psychophysiol.*, vol. 85, no. 1, pp. 125–8, Jul. 2012.

37. B. Penolazzi, C. Spironelli, C. Vio, and A. Angrilli, “Brain plasticity in developmental dyslexia after phonological treatment: a beta EEG band study.,” *Behav. Brain Res.*, vol. 209, no. 1, pp. 179–82, May 2010.