



-RESEARCH ARTICLE-

Emotion Analysis using Different Stimuli with EEG Signals in Emotional Space

Yasar Dasdemir^{1*}, Esen Yildirim², Serdar Yildirim³

¹Iskenderun Technical University, Computer Engineering, Hatay, Turkey

²Adana Science and Technology University, Electric-Electronics Engineering, Adana, Turkey

³Adana Science and Technology University, Computer Engineering, Adana, Turkey

Abstract

Automatic detection for human-machine interfaces of the emotional states of the people is one of the difficult tasks. EEG signals that are very difficult to control by the person are also used in emotion recognition tasks. In this study, emotion analysis and classification study were conducted by using EEG signals for different types of stimuli. The combination of the audio and video information has been shown to be more effective about the classification of positive/negative (high/low) emotion by using wavelet transform from EEG signals, and true positive rate of 81.6% was obtained in valence dimension. Information of audio was found to be more effective than the information of video at classification that is made in arousal dimension, and true positive rate of 73.7% was obtained when both stimuli of audio and audio+video are used. Four class classification performance has also been examined in the space of valence-arousal.

Keywords:

EEG, stimuli types, emotion, emotion space model, valence, arousal

Article history:

Received 07 March 2017, Accepted 05 June 2017, Available online 20 June 2017

* Corresponding Author: Yasar Dasdemir, e-mail: yasar.dasdemir@iste.edu.tr

Introduction

Emotions have a key role in communication and they are required to understand human behavior. Emotion is among the topics of research of several disciplines such as neuroscience, psychology, and linguistics. Especially in the field of psychology, there are different approaches to emotion modeling (Gunes & Pantic, 2010). Categorical and dimensional approaches of these approaches are the most widely used models for labeling emotion in the studies about emotions recognition.

Emotional state in categorical approach has been identified as the mood that is expressed as discrete. Although there is a wide range of defined emotion categories, anger, fear, surprise, happiness, sadness, disgust feeling classes proposed by Ekman et al. (Ekman, 1999) have been accepted as universal emotions. In the dimensional approach, emotions are not limited to a small number of discrete emotion classes, instead of this, it is defined as points in a multi-dimensional space. In this approach, diversity of emotions is considered in 3 dimensions. These dimensions are valence, arousal, and dominance. When valence determines the range from negative to positive of emotions, arousal denotes the range from calmness to exciting of emotions. The dominance dimension is associated with the control of the environment with the feeling.

Dimensional approach is used for the representation and the labeling of the emotions, and studies were performed about valence and arousal dimensions that are widely used in the literature. Effects of emotion analysis and classification of the type of stimuli were also investigated using three different stimuli types as Audio (sound), Video (visual) and Audio + Video (both sound and visual) to reveal the feelings of the participants in the database created under this study.

Literature Review

Many studies have used accepted sense stimulus (movies, pictures, sound, smell) to elicit some emotions. For example, IAPS (International Affective Picture System) (Lang, Bradley, & Cuthbert, 1997) is a dataset that is frequently used in studies on emotion. 716 natural colored images like landscapes, people, and objects which are taken by professional photographers are found inside it, and it is widely used with EEG for emotion recognition studies (Yohanes vd., 2012), (Xu & Plataniotis, 2012), (Ramirez & Vamvakousis, 2012). There are also studies that use movie clips for stimulation of feelings (Rottenberg, Ray, & Gross, 2007).

3-dimensional valence-arousal-dominance or pleasure-arousal-dominances fields are used in cognitive theory. These fields are frequently used in emotion processing studies (Yoon & Chung, 2011), (Liu & Sourina, 2012), (Ahmed, 2014), (Al-galal, Taha, & Wahab, 2015), (Chen & Han, 2015), (Atkinson & Campos, 2016), (Huang vd., 2016).

Ramirez and Vamvakousis have proposed a new method for emotion recognition by using Emotiv EPOC device (Ramirez & Vamvakousis, 2012). They have used some sounds from IADS (International Affective Digitized Sounds) (Bradley & Lang, 1999) sound library which consisted of labeled emotional sounds as stimulators. They tried to classify high/low arousal and high/low valence emotions with various machine learning techniques, using the valence and arousal plane. They did EEG measurements from AF3, AF4, F3 and F4 channels in the section prefrontal cortex. They have used beta/alpha ratio as an arousal status indicator. They have achieved classification performances which are 77.82% for high/low arousal and 80.11% for positive/negative valence. The performance values obtained by SVM with radial basis function kernel classifier.

The achievement was obtained 64.84% and 61.17% ratio about the classification of high/low arousal and valence of emotions, respectively. The experiment is made with bispectrum analysis of EEG signals by using valence-arousal emotion space (Kumar, Khaund, & Hazarika, 2016).

Emotion recognition studies have been conducted on DEAP dataset (Koelstra vd., 2012) obtained by collecting video stimuli and EEG records using wavelet-based attributes (Srinivas, Rama, & Rao, 2016) and emotion characteristics emerged in the delta, theta, alpha, beta and gamma band are classified with MLP (Multi-Layer Perceptron) and RBF (Radial Basis Function). The best results in the frequency field have been obtained with 54.54% of the RBF and 63.63% of the MLP.

Emotion analysis and classification studies are made in this study using EEG signals, and performance of the system was tested in the 4-quadrant emotion of field. Necessity database has been established within this scope to perform studies. As the database is created, emotion targeted using only audio, only video, and audio+video stimuli have been triggered. Short videos obtained from domestic and foreign films were used as stimuli within this scope. Each video is selected as 60-second segments, and 3 different versions of the video (only audio, only video, and both audio and video) are used as stimuli.

Background

EEG signals and facial expressions collected from 25 volunteers are used in this study. In the database (Duygu-DB) created in the study, EEG signals were recorded using an Emotive EPOC wireless EEG device and the facial expression videos were recorded using a smartphone with 1920x1080 HD 30 fps resolution. The videos are not used in this study. 15 different piece of film containing different mood with 60 seconds length is used as stimuli. Data were collected in a single session presenting stimuli to the participants in 3 different ways (only audio, only video, audio+video). During the collection of data, stimulants were presented to each participant in a random order. Obtained Duygu-DB database's content is given in Table 1.

Table 1. Duygu-DB content summary

Subjective annotation	
Number of videos	45 (15 audio only, 15 videos only, 15 audio+video)
Video duration	1 minute
Stimuli selection method	130 manually selected via YouTube and movies, 15 as for that 11 participants (8 male, 3 female) selected that didn't attend experiment from 130 video (15 of stimulus are the most distance from origin, which they are selected by 11 subjects)
Number of ratings per video	6
Rating scales	Valence, Arousal, Dominance
Rating values	Discrete scale of 1-9 (0-90 limit values)
Physiological Experiment	
Number of participants	25 (20 male, 5 female; age between 18 and 27) Average age: 20.52, standard deviation: 1.69
Number of videos	1125 (45x25)
Recorded signals	16-channel 128Hz EEG Face videos (front face, 1920x1080 HD 30 fps smartphone)

Self-Assessment Form

Self-Assessment Form is used to determine the status of participants' mood that stimulus are evoked in themselves. The personal assessment measures are often used for emotion analysis in research on emotion. SAM (Self-Assessment Manikin) images (Morris, 1995) are used for making these assessments (Figure 1). SAM visuals use the 3D-emotion field that is represented a feeling with Valence, Arousal and Dominance values.

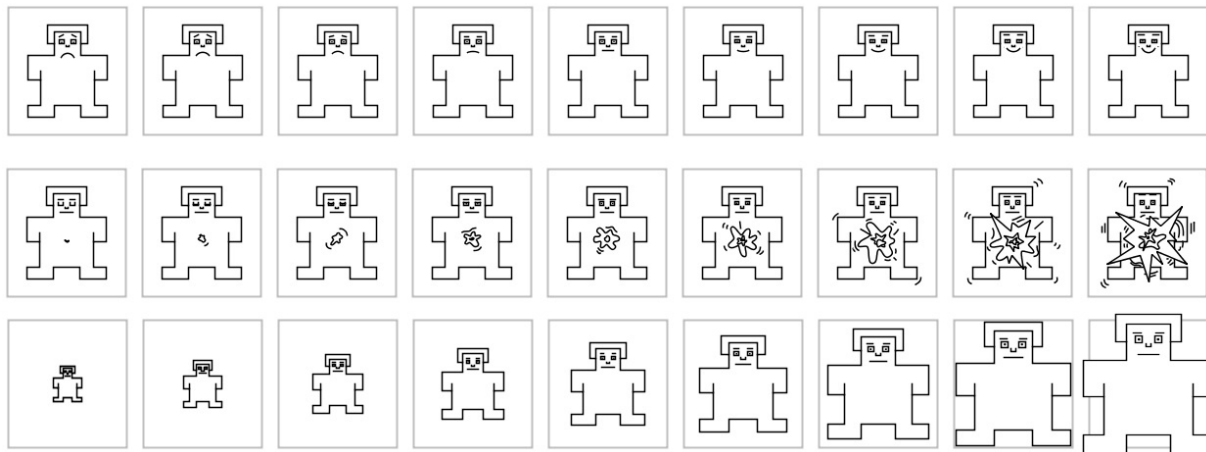


Figure 1. The Self-Assessment Manikin (SAM) visuals (top panel-valence, middle panel-arousal, bottom panel-dominance)

In the figure of valence, expressions such as unhappy, sad, frustrated, desperate and discontent that are totally corresponding to negative values are located on the left, expressions such as happy, contented, pleasant, optimistic and satisfaction that are corresponding to positive values are located on the right.

The visual of arousal measures the severity and intensity of emotion. In this picture, while expressions such as very quiet, peaceful, numb, unstimulated are located on the left, the expressions such as crazy, exciting, passionate and stimulated are located on the right.

While the expression that is under the influence of other, reckless, dispirited, unstable is located on the left of the scale of dominance, the expression which is quite stable, which know what it is doing is located on the right of the scale of dominance. The left side of the image indicates that the event is not under the control of the person, but the right side of the image indicates that the events are under control of the person. Dominance dimension is also used to distinguish angry feelings from the scare. Dominance dimension is needed for making the distinction between specifically similar valence and arousal values of emotions.

Selection of Stimulus

Emotional content film pieces were used as stimuli in 3 ways in the experiment. First, audios that are belonging to the part, secondly, videos that are belonging to the part and thirdly both audio and video that are belonging to the part are used together. These 3 different stimuli have been applied to participants in random order.

A minute's 130 stimuli were selected from firstly 59 films (9 foreign, 50 domestic) for selecting of the stimuli that will be used in the experiment by using film reviews on the internet. While

selecting, movie clips were selected according to evaluators' ratings such that they are distributed equally to 4 regions in 2-dimensional Valence/Arousal space.

Movie clips have been watched to 11 evaluators (8 male, 3 female, the average of age: 20.91, standard deviation: 1.44) who don't participate in the experiments of data collection, and selecting appropriate intervals through the SAM images was requested.

Evaluation of discrete emotion class in addition to evaluation of emotions dimension was also requested from the evaluators. Used emotion discrete classes were determined as happy, surprise, disgust, frustrated, scared, bored, sad, satisfied, calm, neutral (natural).

Movie clips that will be used in data collection are selected with equally distributed to the 4 areas of valence-arousal in two-dimensional yield by considering the value of the chosen evaluator. The emotional highlight value is used to make this choice (Koelstra at al., 2012). Emotional stress value for the clips, e_i , is the distance of the emotional content of the movie clip to the origin, and is calculated as in (1).

$$e_i = \sqrt{a_i^2 + v_i^2} \quad (1)$$

Where a_i shows the arousal value of the clip i ., and v_i shows the valence value of the clip i . Low e_i values (close to the original values) will show the proximity to the neutral mood. 15 clips at the extreme point of the e_i values representative of 4 emotion areas were selected in this study.

Stimulus are watched at a single session in random order according to the experimental protocol prepared to the participants. Meanwhile, face video and EEG signals were recorded. After each stimulus, participants filled the self-assessment form. Synchronization of collected data has been achieved by following the time of filling form within the time of the relaxation.

Data Acquisition

Multichannel EEG data used in the experiment were collected by Emotiv EPOC that is cheap, light and wireless BCI (Brain-Computer Interface) device. Emotiv EPOC EEG device uses 14 number of electrodes which are AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, O2.

25 participants who composed of associate degree students between 18 and 27 years of age (20 male, 5 female, average age: 20.52, standard deviation: 1.69) participated in the data collection process voluntarily.

EEG recordings were obtained with the approval, considering ethical principles of World Medical Association Declaration of Helsinki related to medical research that is done on humans by Mustafa Kemal University Faculty of Medicine. Informing participants about this subject, their approval was obtained with confirmation (consent) form. One day ahead, experiment protocol described to participants, and they have been warned that they must have enough sleep and do not take any stimulant and come with an empty stomach on the experiment day.

Stimuli are presented to the participants in random order according to the following steps:

1. Short beep sound as the first stimuli,
2. Watching 60 seconds of stimuli (3 different stimuli are shown in random order),
3. 30 seconds for filling the self-assessments form,
4. 10 seconds for relaxation stage (black screen).

Material and Method

In this study, followed steps are shown in Figure 2, the methods that are used are given below.

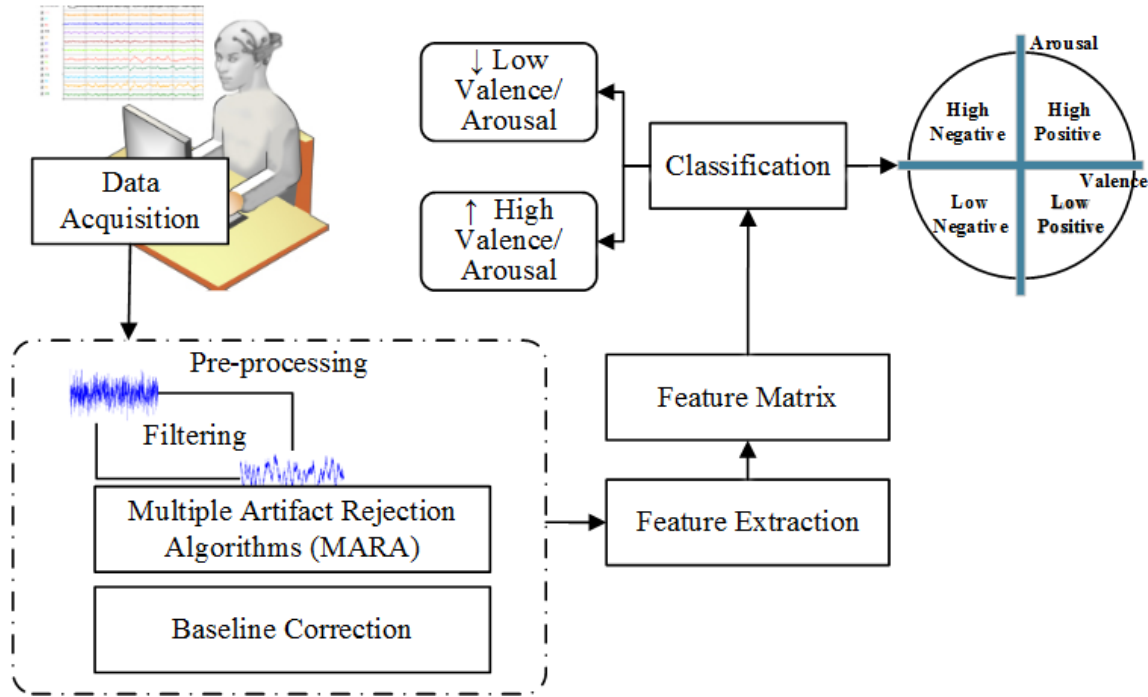


Figure 2. The flow chart of the study

Pre-processing

EEG device that is used was not effected so much from the negativity coming from power line because of the fact that it provides wireless. Third order Butterworth band-pass filter with 0.16 and 43 Hz cut-off frequencies was applied to the collected data. Artifacts such as EOG (electrooculogram) and EMG (electromyogram) can be cleaned with this filtration. After that, independent components were cleaned with MARA (Multiple Artifact Rejection Algorithms) Tool in Matlab software (Winkler, Haufe, & Tangermann, 2011). 10 sec. reference level (baseline) were removed from the signal during the average recording of the stimuli in addition to correct stimuli-unrelated variations. Making this allows us to achieve the changing of the EEG data that occurred during the stimuli.

Feature Extraction and Feature Matrix

Discrete Wavelet Transform that is a suitable method is used up to the fourth level for the analysis of non-stationary signals like EEG signal about the extract of attribution. In this study, EEG Signals are separated to the 5 different frequency bands by using “db4” wavelet function. The frequency bands that is used are given in Table 2.

Table 2. The levels of wavelet analysis (A: Approximation coefficients, D: Detail coefficients)

Decomposition levels	Frequency range (Hz)	EEG frequency bands
D1	32 – 64	Gamma
D2	16 – 32	Beta
D3	8 – 16	Alpha
D4	4 – 8	Theta
A4	0 - 4	Delta

Attributes comprise of the maximum and minimum values, standard deviations, average value, exchange value, average power values and the entropies of the obtained coefficients at each decomposing level. In total, 295 units of attributes were extracted for each sample.

Binary Classification from EEG Signals

Studies in this field, high/low classification of valence, arousal, and dominance of emotion dimension is made from EEG signals that are collected by using different stimuli. Impacts on the classification performance of different types of stimuli were compared in the classification studies.

Table 3 shows the effects of stimuli types to the valence, arousal and dominance dimensions. Values in the table are the true positive rate that is obtained by 10-fold cross-validation by using MLP and RF (Random Forest).

Table 3. Various classifier results for different stimulus types

Classifiers	Stimuli types	Valence	Arousal
MLP	Audio	0.763	0.658
	Video	0.711	0.658
	Audio+Video	0.816	0.684
RF	Audio	0.659	0.737
	Video	0.686	0.684
	Audio+Video	0.730	0.737

It was obtained for the best performance audio+video stimuli in terms of true positive rates at high/low classification made in valence dimension. These results show that using both of the audio and video information is more effective about the discrimination of negative/positive emotions. Highest performance was obtained using MLP classifier with 81.6%.

The best performance in the binary classification that is made in arousal dimension was obtained using RF classifier with 73.7%. When the results were analyzed, it was observed that only image information showed the lower performance than just sound knowledge about the classification, and audio information is more effective about the discrimination from calmness to exciting.

In Table 4, accuracy, sensitivity and f-score values for different types of stimuli are given for each classifier and each class.

Table 4. Precision, Recall and f-score values of classifiers for different stimulus types

		MLP				RF			
		Valence		Arousal		Valence		Arousal	
		Low	High	Low	High	Low	High	Low	High
Audio	Precision	0.750	0.833	0.710	0.429	0.781	0.833	0.722	1.000
	Recall	0.960	0.385	0.846	0.250	0.962	0.417	1.000	0.167
	f-score	0.842	0.526	0.772	0.316	0.757	0.429	0.839	0.286
Video	Precision	0.808	0.500	0.783	0.467	0.808	0.333	0.688	0.667
	Recall	0.778	0.546	0.692	0.583	0.778	0.375	0.917	0.286
	f-score	0.793	0.522	0.735	0.519	0.792	0.353	0.786	0.400
Audio + Video	Precision	0.864	0.750	0.692	0.680	0.759	0.625	0.813	0.681
	Recall	0.826	0.800	0.529	0.809	0.880	0.417	0.650	0.833
	f-score	0.844	0.774	0.600	0.739	0.815	0.500	0.722	0.750

Four-Region Classification from EEG Signals

In this chapter, the four classification performances of the EEG signal were investigated for four-zone valence-arousal space (Figure 3).

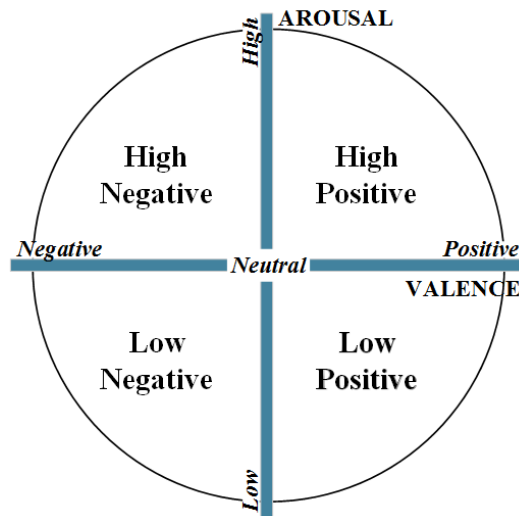


Figure 3. The delineation of the four-region emotion space

Values in Table 5 are the true positive rates that are obtained by 10-fold cross-validation by using MLP, RF, and SVM (Support Vector Machines) classifiers. It is seen that the success obtained by sound stimuli from the EEG channel thanks to the values. Performances that is close each other were obtained with only MLP classifier just for audio and audio+video stimuli. The low success from the EEG channel was obtained just with the video stimuli. It was observed that using only videos as the stimuli showed lower performance about the zoning of the emotions compared with other stimuli types. The highest performance was obtained with the sound stimuli.

Table 5. Classifier performances

Stimuli Types	MLP	RF	SVM
Audio	0.411	0.507	0.437
Video	0.373	0.440	0.395
Audio+Video	0.419	0.467	0.419

Weighted average F-score values belong to the classifier are given in Table 6. As can be seen overall from the results, it is seen that better overall classification performance is obtained with the audio stimuli.

Table 6. Weighted average f-score values

Stimuli Types	MLP	RF	SVM
Audio	0.381	0.498	0.398
Video	0.370	0.385	0.389
Audio+Video	0.363	0.392	0.398

Result and Discussion

In this study, emotion analysis and classification study are made for different types of stimuli using EEG signals. While EEG signals were collecting, three different type of stimuli as audio, video, audio+video were used for stimuli that are demanded from participants. Dimensional approach is used for the representation and the labeling of the emotions, and studies about valence and arousal dimensions that is widely used in the literature are made.

Two different classifiers were used for wavelet transform and binary classification about attribute extraction from EEG signals. The results were compared in the studies of classification in valence and arousal dimensions with audio, video and audio+video stimuli. It is shown that using both of the audio and video information is more effective about the binary classification of high/low that is made in valence dimension. It is observed that using audio information is more effective about the binary classification of high/low that is made in arousal dimension.

Classification performance was presented in the four classification study again with attributes that are obtained by the wavelet transform EEG channel. It is seen that the success obtained by the EEG channel sound stimuli is higher in general. The lowest performance has been achieved by only the video stimuli at the classification studies that is made by using EEG. The highest performance was obtained with the audio stimuli.

In future studies, the use of attributes to be obtained in time and frequency dimensions and the effect of attribute selection on classification performance can be examined.

References

- Ahmed, M. A. (2014). Emotion Recognition Based on Correlation Between Left and Right Frontal EEG Assymetry. *Mecatronics (MECATRONICS)*, 99–103.
- Al-galal, S. A. Y., Taha, I. F., & Wahab, A. (2015). Relaxing Music Using Valence-Arousal Model. *Içinde 4th International Conference on Advanced Computer Science Applications and Technologies* (ss. 9–14).

- Atkinson, J., & Campos, D. (2016). Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers R. *Expert Systems With Applications*, 47, 35–41. <http://doi.org/10.1016/j.eswa.2015.10.049>
- Bradley, M. M., & Lang, P. J. (1999). *International Affective Digitized Sounds (IADS): Stimuli, Instruction Manual and Affective Ratings*. Technical Report B-2, Gainesville, FL: The Center for Research in Psychophysiology, University of Florida, FL, USA.
- Chen, M., & Han, J. (2015). Identifying Valence and Arousal Levels via Connectivity between EEG Channels. İçinde *International Conference on Affective Computing and Intelligent Interaction (ACII)* (ss. 63–69).
- Ekman, P. (1999). Basic Emotions. İçinde *Handbook of Cognition and Emotion* (Dalglish, ss. 45–60). New York: John Wiley&Sons Ltd.
- Gunes, H., & Pantic, M. (2010). Automatic , Dimensional and Continuous Emotion recognition. *International Journal of Synthetic Emotions*, 1(1), 68–99.
- Huang, X., Kortelainen, J., Zhao, G., Li, X., Moilanen, A., Seppänen, T., & Pietikäinen, M. (2016). Multi-modal emotion analysis from facial expressions and electroencephalogram. *Computer Vision and Image Understanding*, 147, 114–124.
- Koelstra, S., Mühl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... Patras, I. (2012). DEAP: A database for emotion analysis; Using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), 18–31.
- Kumar, N., Khaund, K., & Hazarika, S. M. (2016). Bispectral Analysis of EEG for Emotion Recognition. *Procedia - Procedia Computer Science*, 84, 31–35. <http://doi.org/10.1016/j.procs.2016.04.062>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings*. University of Florida.
- Liu, Y., & Sourina, O. (2012). EEG-based Valence Level Recognition for Real-Time Applications. İçinde *International Conference on Cyberworlds* (ss. 53–60).
- Morris, J. D. (1995). Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. *Journal of advertising research*, 35(6), 63–68.
- Ramirez, R., & Vamvakousis, Z. (2012). Detecting Emotion from EEG Signals Using the Emotive Epoc Device. İçinde *Brain Informatics* (ss. 175–184).
- Rottenberg, J., Ray, R. D., & Gross, J. J. (2007). Emotion elicitation using films. İçinde *The handbook of emotion elicitation and assessment* (ss. 9–28). Oxford University Press.
- Srinivas, V., Rama, V., & Rao, C. B. R. (2016). Wavelet Based Emotion Recognition Using RBF Algorithm. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, 4(5), 29–34.
- Winkler, I., Haufe, S., & Tangermann, M. (2011). Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals. *Behavioral and Brain Functions*, 7(1), 30.
- Xu, H., & Plataniotis, K. N. K. (2012). Affect Recognition Using EEG Signal. İçinde *MMSP 2012* (ss. 299–304).
- Yohanes, R. E. J., Member, S., Ser, W., Member, S., Huang, G., & Member, S. (2012). Discrete Wavelet Transform Coefficients for Emotion Recognition from EEG Signals. İçinde *34th Annual International Conference of the IEEE EMBS* (ss. 2251–2254).
- Yoon, H. J., & Chung, S. Y. (2011). EEG Spectral Analysis in Valence and Arousal Dimensions of Emotion. İçinde *11th International Conference on Control, Automation and Systems* (ss. 1319–1322).