Emotion Recognition via Multi Channel EEG Signal Fusion and Pattern Recognition

Deger Ayata[1], Yusuf Yaslan[1], Mustafa Kamaşak[1]
[1]Computer Engineering Faculty, Istanbul Technical University, Istanbul, Turkey
{ayatadeger, yyaslan, kamasak}@itu.edu.tr

***Abstract*— Compared with facial expression based approaches, EEG is a reliable approach to probe the internal cognitive and emotional changes of users. EEGs high temporal resolution provides large amount of useful information in analyzing the humans real emotional state directly. We have studied emotion recognition from multi-channel EEG signals using wavelet transform, time based features and data fusion techniques. We have categorized valence and arousal and studied relationship between EEG signals and arousal and valence using Random Forests and k-Nearest Neighbor algorithms. We have obtained 82.18 % and 82.98 % accuracy rate for arousal and valence respectively.**

Keywords — Emotion Recognition, Pattern Recognition, Data Fusion, Feature Level Fusion, Multi-channel sensor processing, Biomedical Signal Processing, EEG, kNN, Random Forests

I. INTRODUCTION

Emotions play a significant and powerful role in everyday life of human beings. The importance of emotions in human-human interaction provided the basis for researchers in the computer engineering, artificial intelligence (AI), biomedical and electronics engineering communities to develop automatic ways for computers to recognize emotional expression, as a goal towards achieving human-computer intelligent interaction [1]. Facial expressions have been used for emotion detection. Ekman et al. stated that six different facial expressions (fearful, angry, sad, disgust, happy, and surprise) were categorically recognized by humans from distinct cultures using a standardized stimulus set [2]. However, using only facial expression signals has disadvantages: using only them is not reliable to detect emotion, especially when people want to conceal their feelings. In recent years, more and more researchers have started to use electroencephalogram (EEG) signals in recognizing emotion because they are reliable. Compared with facial expression, EEG is a reliable approach to probe the internal cognitive and emotional changes of users. EEGs high temporal resolution provides large amount of useful information in analyzing the humans real emotional state directly [3].

In this study, emotion recognition from EEG signals was performed using discrete wavelet transform and time based statistical features. Data fusion techniques have been used to combine data from 32 different EEG source channels and feature level fusion has been implemented.

The emotion valence-arousal dimensional model has been used to measure degree of attraction of a person toward selected youtube videos. Valence and arousal have been categorized and the relationship between EEG signals and arousal and valence has been studied using machine learning algorithms such as Random Forests and k-Nearest Neighbour (kNN).

II. RELATED WORK

EEG measures the summation of electrical activity on the scalp; primarily derived from post-synaptic activity around the dendrites of pyramidal neurons in the cerebral cortex [3,4,5,6,7,12].

There exist extensive research studies in EEG based emotion recognition domain. Choppin confirmed that EEG signals can be used for emotion recognition and got a classification accuracy of 64 % based on neural networks. Choppin has built an emotion recognizer for patients suffering from amyotrophic lateral sclerosis [4]. These patients lose, at some moment during their disease, their ability to use their muscles, and additionally they lose their ability to speak or to communicate on any other way. For that reason, Choppin uses only EEG signals to recognize emotion, in order to provide those people with the possibility to express their feelings. He uses neural networks to classify the EEG signals online, and achieved a correct classification rate for new unseen samples of about 64 %, when using three emotion classes. Channel et al. has conducted a research on emotion assessment related to arousal evaluation using EEG's and peripheral physiological signals. They have used Galvanic Skin Resistance (GSR), blood pressure, temperature as well as EEG data. They have reported that EEG can be used in arousal recognition. They have used Naïve Bayes and Fisher Discriminant Analysis (FDA) classifiers. They have also reported that when fusing EEG and peripheral features, the improvement was better with FDA than with the Bayes classifier [5]. Murugappan et al. has worked on EEG Feature Extraction for Classifying Emotions using Fuzzy C-Means (FCM) and Fuzzy K–Means (FKM). They have collected EEG data from 6 subjects with in an age group of 21-27 using 63 biosensors. After preprocessing the signals, discrete wavelet transform is employed to extract the EEG parameters [6]. Results confirm the possibility of using wavelet transform based feature extraction for assessing the human emotions from EEG signals. Lin et al. has worked on EEG-based emotion recognition in music listening. They applied machine–learning algorithms to categorize EEG dynamics according to subject self- reported emotional states during music listening activity [7]. They used SVM to classify four emotional states (pleasure, joy, anger and sadness) and obtained an average classification accuracy of approx. 83 %. Nie et al. studied the relationship between EEG signals and human emotions [8]. Nie has used EEG signals to classify emotions either positive or negative. Support Vector Machine (SVM) has been used for classification. They reported 87.53 % of accuracy between positive and negative emotions. Wang et al. presented a new method based on a wrapped Sparse Group Lasso for channel and feature selection of fused EEG signals [12].

In our previous works, we have studied emotion recognition only from Galvanic Skin Response (GSR) signals [13, 14]. In this study, we are focusing on multichannel EEG signals using feature fusion. With the current study, we identify the user's current emotional state by predicting arousal and valence values while the participant is listening to music. The proposed method can be used in emotion aware music therapy or music recommendation engines applications that consider the participants emotion states and needs.

III. MATERIALS AND METHODS

A. EEG SIGNALS

The intensities of brain waves recorded from the surface of the scalp range from 0 to 200 microvolts, and their frequencies range from once every few seconds to 50 or more per second. A typical adult human EEG signal is about 10μV to 100 μV in amplitude when measured from the scalp and is about 10–20 mV when measured from subdural electrodes. EEG has low spatial resolution and is non-stationary in nature. EEG signals are not periodic and their amplitude, phase and frequencies change.

There are mainly five types of Brain waves: Delta waves (0.4-4 Hz), which occurs in sleeping adults, premature babies or if there is any sub cortical lesions and is found in the frontal region of brain in adults and posterior region in children , Theta waves (4-8 Hz) which occurs in children, in adults when they are in emotional stress or they have deep midline disorders and is found in parietal and occipital region, Alpha waves (8-13 Hz) which occurs in quiet resting state but not sleep and is found in the occipital region, Beta waves(13-30 Hz) which occurs in active, busy concentration or anxious thinking state and is found in the frontal and parietal region and Gamma waves (30-100 Hz) and occurs in certain cognitive or motor functions. Brain generates more than one brainwave pattern at a time. Brain is seldom in “just” beta, or “just” Alpha, or “just” any one brain wave state. Brain fluctuates between states, for example emitting beta waves and theta waves simultaneously, with one of them being predominant.

B. EMOTION REPRESENTATION

The emotion valence-arousal dimensional model, represented in Figure 1, is widely used in many research studies. The Pleasure - Displeasure Scale measures how pleasant an emotion may be. Pleasure (Valence) ranges from unpleasant to pleasant and it is the degree of attraction of a person toward a specific object or event. It ranges from negative to positive. The Arousal-Non Arousal Scale measures the intensity of the emotion. The arousal is a physiological and psychological state of being awake or reactive to stimuli, ranging from passive to active. Valence-arousal model chart, is a model for emotions to be mapped out by range of arousal and valence that is experienced during a particular emotion. The Valence - Arousal-axis separate the coordinate plane into four regions. If arousal>0 and valence>0 then a person can be in one of {Happy, Excited, Pleased} emotion states.

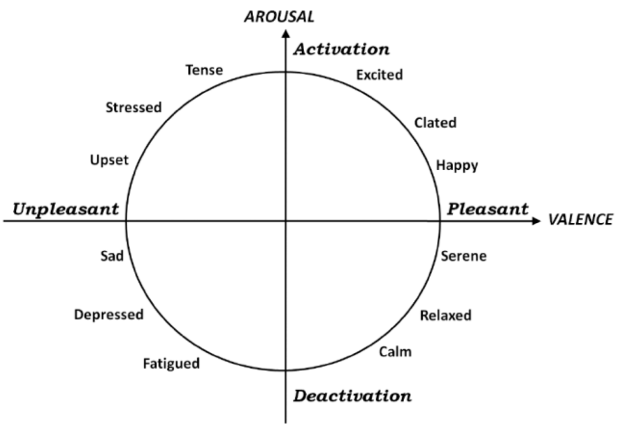

Figure 1. Valence - Arousal Model

C. DATASET

Deap is a multimodal dataset for the analysis of human affective states, in the dataset EEG and peripheral physiological signals of 32 participants were recorded as each watched 40 videos, each video is one-minute long excerpts of music videos [9]. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance and familiarity. For 22 of the 32 participants, frontal face video was also recorded. The data was down sampled to 128Hz, EOG artefacts were removed, a bandpass frequency filter from 4.0 - 45.0 Hz was applied and, the data was segmented into 60 second trials and a 3 second pre-trial. The total signal record time for each video is 63 second and sampling frequency is 128 Hz which means for each channel 8064 sample data points have point collected. For each video 32 EEG channels data have been collected and sampled with 128Hz. Table 1 depicts channel no and channel content mapping. The table the EEG channel names (according to the 10/20 system) for both locations and the indices that can be used to convert one ordering to the other. Each site has a letter to identify the lobe and a number to identify the hemisphere location. The letters F, T, C, P and O stand for frontal, temporal, central, parietal, and occipital lobes, respectively. (Note that there exists no central lobe; the "C" letter is used only for identification purposes.) Even numbers (2,4,6,8) refer to electrode positions on the right hemisphere, whereas odd numbers (1,3,5,7) refer to those on the left hemisphere. A "z" (zero) refers to an electrode placed on the midline.

D. FEATURE EXTRACTION

Features from EEG signals have been extracted based on the time, frequency attributes and statistics. All 32 EEG channel signals have been used.

Time Domain Features

In each channel, the corresponding EEG signals have been subjected to 3 seconds length moving windows for feature extraction. In each trial, we have obtained 32 channels’ signals and divide each channel signal into 20 segments with 3s length per segment. EEG features have been first extracted from each window, and their values across the consecutive windows have been concatenated for each subject and for each video. In the time domain, arithmetic mean value, maximum value, minimum value, standard deviation, variance, skewness coefficient, kurtosis coefficient, median, number of zero crossings, entropy, mean energy value have been considered as features. Since we have 20 segments (20 segments, each 3 seconds length and in total 1 minute) we have 220 extracted time domain statistics (20*11 features) for each signal.

Table 1: EEG Channels Data

Channel#	Channel	Channel#	Channel
1	Fp1	2	AF3
3	F3	4	F7
5	FC5	6	FC1
7	C3	8	T7
9	CP5	10	CP1
11	P3	12	P7
13	P03	14	O1
15	Oz	16	Pz
17	Fp2	18	AF4
19	Fz	20	F4
21	F8	22	FC6
23	FC2	24	Cz
25	C4	26	T8
27	CP6	28	CP2
29	P4	30	P8
31	PO4	32	O2

Discrete Wavelet Transformation Features

EEG has low spatial resolution and is non-stationary in nature. EEG signals are not periodic and their amplitude, phase and frequencies change. Discrete Wavelet Transformation is a method developed to overcome the deficiencies of the Fourier transformation over non-stationary signals and this method is less sensitive towards noise and can be easily applied to non-stationary signals. Features have been extracted from the EEG spectrum calculated using Discrete Wavelet Transform. For the DWT, it is important to identify appropriate wavelet type and determining the level of decomposition. Daubechies db2 and db4 has been used and compared as wavelet.

The features are the sum of absolute amplitudes, min, max, mean energy, sum of squares, kurtosis, skewness, standard deviation in the full, Theta Band, Alpha Band, Beta Band and Gamma Band. Since we have 8 features for each band, in total we have extracted 5*8 features, 40 in total.

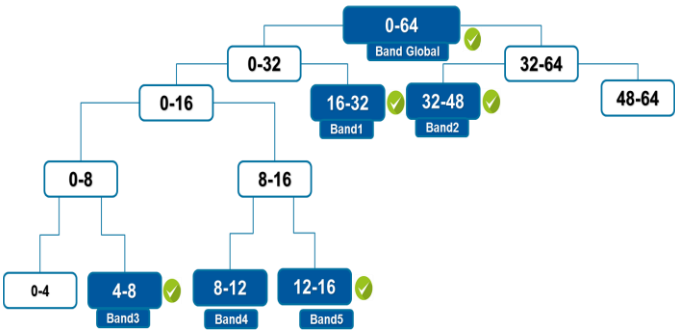

Figure 2. DWT- Used Bands

E. FEATURE FUSION

Main objective of employing fusion is to produce a fused result that provides the most detailed and reliable information possible. Fusing multiple information sources together also produces a more efficient representation of the data. Data fusion is categorized as feature level and decision level. Feature level fusion requires the extraction of different features from the source data – before features are merged together. A feature-level fusion scheme integrates unimodal features before learning concepts. The main advantage of this scheme is the use of only one learning stage and taking advantage of mutual information from data. Decision level fusion combines the results from multiple algorithms to yield a final fused decision. Decision-level fusion, or fusion of classifiers, consists of processing the classification results of prior classification stages. The main goal of this procedure is to take advantage of the redundancy of a set of independent classifiers to achieve higher robustness by combining their results [11]. We have used feature-level fusion. The feature vectors selected from all 32 channels. Each EEG signal has been segmented into 20 windows with 3 seconds length and time domain statistics have been obtained, but also for each full EEG signal also global features have been extracted. Feature vectors have been used directly and by concatenation into RF and kNN classifiers.

F. CLASSIFICATION

Labeling the samples is critical for Machine Learning. arousal and valence values have been categorized to two (Low, High) classes. We divide the trials into classes according to each trial's rating value (high: ≥ 4.5, low: < 4.5). 32 EEG signals taken from 32 subjects all have been used for training and test steps.

After feature extraction the signals are classified into various classes using Random Forest and K-Nearest Neighbours. kNN is a family of simple classification and regression algorithms. Random Forests are an ensemble method with which classification and regression are performed using a forest of decision trees, each constructed using a random subset of the features [10].

IV. EXPERIMENTAL RESULTS

10-fold cross validation has been used to evaluate the accuracy rate. Test has been conducted with 10-fold cross validation by using Random Forest and kNN machine learning algorithms. Tests are conducted for time domain statistics features, wavelet transformation features and fusion of time domain and wavelet coefficients each separately. Each extracted and fused feature set has been tested with KNN and Random Forest separately.

Time Domain Statistics Based Experiments

Time Domain Statistics results are depicted in Table 2. kNN is has higher recognition rate compared to Random Forests. We have 32 subjects (participants) dataset. Each subject has watched 40 video (each video is 1 minute). For each video 32 EEG channels data has been used for feature extraction. Each EEG signal has been segmented into 20 windows with 3 seconds length and time domain statistics has been obtained, but also for each full EEG signal also global features has been extracted. Feature vectors have been used directly and by concatenation into classifiers.

Table 2. Time Domain Results – 2 Class (Low, High)

Classifier	Training Instance#	Arousal Acc(%)	Valence Acc(%)
RF	32*40*32	78.19	78.3
RF	32 * 40	70.83	70.43
kNN	32*40*32	82.18	82.98
kNN	32 * 40	64.18	65.46

Discrete Wavelet Transformation Based Experiments

DWT based results are depicted in Table 4. DWT has been applied to each EEG signal separately and 40 features has been extracted from 5 bands shown in Figure 2. (8 features per band)

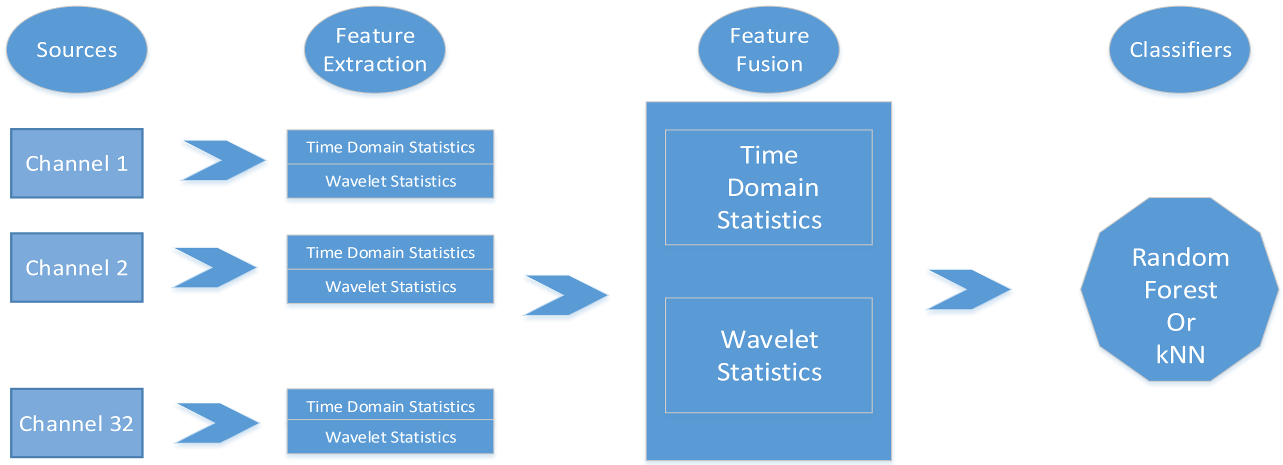

Figure 3. Feature Fusion and Classification Pipeline

Table 3. Discrete Wavelet Results – 2 Class (Low, High)

Classifier	Wavelet	Arousal Acc(%)	Valence Acc(%)
RF	db2	80.05	77.34
RF	db4	80.12	71.43
kNN	db2	74.60	80.82
kNN	db4	75.54	80.74

Feature Fusion: Time Domain and Discrete Wavelet Features Experiments

Extracted time domain and DWT features are combined and resulted feature vector is fed into RF and kNN classifiers, results are depicted in Table 4 .

Table 4. Feature Fusion Results - 2 Class (Low, High)

Classifier	Wavelet	Arousal Acc(%)	Valence Acc(%)
RF	db2	70.91	70.39
RF	db4	70.51	71.64
kNN	db2	63.38	63.82
kNN	db4	65.56	63.90

V. CONCLUSION AND FUTURE WORK

We have studied emotion recognition from multi-channel EEG signals using wavelet transform, time based features and data fusion techniques. We have categorized valence and arousal and studied relationship between EEG signals and arousal and valence using Random Forests and k-Nearest Neighbor algorithms. Extracted features capture the emotional changes of the subject through their EEG signals. EEGs high temporal resolution provides large amount of useful information in analyzing the human real emotional state directly. This study shows a significant relationship between EEG signals and emotional states experienced by the subjects during the interaction with visual-audio content. We have obtained 82.18% and 82.98% accuracy rate for arousal and valence respectively.

Extracted features number will be enriched and physiological including signals Galvanic Skin Response, Respiration Belt, Plethysmograph and Temperature will be combined with EEG signals using data fusion and ensemble methods to increase accuracy rate. In the future, we are interested in unsupervised feature learning for improving accuracy. We are also planning to conduct more experiment data and support from clinical EEG paradigm in the future work.

REFERENCES

[1] N. Sebe, I.Cohen, and T. S. Huang, "Multimodal Emotion Recognition", WSPC, June 18, 2004

[2]P. Ekman, P., R.W.Levenson, , W.V. Friesen. Autonomic nervous system activity distinguishing among emotions. Science 221, 1208– 1210., 1983

[3] L. Xiang, P. Zhang P, Song D, Yu G, Hou Y, Hu B. EEG Based Emotion Identification Using Unsupervised Deep Feature Learning. 2015. NeuroIR'15.

[4] A. Choppin, "EEG-based human interface for disabled individuals: emotion expression with neural network," Thesis, Tokyo Institute of Technology, 2000.

[5] G. Channel, J. Kronegg, D. Grandjean, T. Pun, "Emotion Assessment : Arousal Evaluation Using EEG's and Peripheral Physiological Signals", Technical Report, Universite de Geneve, 2005.

[6] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry , "EEG Feature Extraction for Classifying Emotions using FCM and FKM", International Journal of Computers and Communications, Issue 2, Volume 1, 2007

[7] Y. Lin, C. Wang, T. Jung, T. Wu, "EEG –based Emotion Recognition in Music Listening", IEEE Transactions on Biomedical Engineering, Vol 57, Issue 7, 2010.

[8] D. Nie, X.Wang, L.Shi, B. Lu, "EEG-based emotion recognition during watching movies," Proceedings of the 5th international IEEE EMBS conference on Neural Engineering, pp. 667-670, 2011.

[9] S. Koelstra, C. Muehl, M. Soleymani, A. Yazdani, J.-S. Lee, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A Database for Emotion Analysis Using Physiological Signals", IEEE Trans. Affective Computing, vol. 3, no. 1, pp. , Jan- Mar. 2012.

[10] Ho, Tin Kam. Random Decision Forests (PDF). Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.

[11] S.Planet, I. Iriondo, "Comparison between Decision-Level and Feature-Level Fusion of Acoustic and Linguistic Features for Spontaneous Emotion Recognition", Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on. IEEE, 2012.

[12] J. Wang, F. Xue, H. Simultaneous, "Channel and Feature Selection of Fused EEG Features Based on Sparse Group Lasso", BioMed ResearchInternational, Volume 2015 (2015), Article ID 703768

[13] D.Ayata, Y.Yaslan, M.Kamasak (2017), "Emotion Recognition via Galvanic Skin Response: Comparison of Machine Learning Algorithms and Feature Extraction Methods". Istanbul University - Journal of Electrical & Electronics Engineering, 17 (1): 3129-3136

[14] D. Ayata, Y. Yaslan ,M. Kamaşak, "Emotion recognition via random forest and galvanic skin response: Comparison of time based feature sets, window sizes and wavelet approaches," 2016 Medical Technologies National Congress (TIPTEKNO), Antalya, 2016, pp. 1-4. doi: 10.1109/TIPTEKNO.2016.7863130