Emotion Recognition via Random Forest and Galvanic Skin Response: Comparison of Time Based Feature Sets, Window Sizes and Wavelet Approaches

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

***Abstract* - Emotions play a significant and powerful role in everyday life of human beings. Developing algorithms for computers to recognize emotional expression is a widely studied area. In this study, emotion recognition from Galvanic signals was performed using time domain and wavelet based features. Feature extraction has been done with various feature set attributes. Various length windows have been used for feature extraction. Various feature attribute sets have been implemented. Valence and arousal have been categorized and relationship between physiological signals and arousal and valence has been studied using Random Forest machine learning algorithm. We have achieved 71.53% and 71.04% accuracy rate for arousal and valence respectively by using only galvanic skin response signal. We have also showed that using convolution has positive affect on accuracy rate compared to non-overlapping window based feature extraction.**

Keywords — Biomedical Signal Processing, Emotion Recognition, Pattern Recognition, Machine Learning, Physiological Signal, Galvanic Skin Response, Random Forest

I. INTRODUCTION

Emotions play a significant and powerful role in everyday life of human beings. The importance of emotions motivated the researchers in the biomedical engineering, computer and electronics engineering disciplines to develop automatic methods for computers to recognize emotional expressions [1]. For a rich set of applications including human-robot interaction, computer aided tutoring, emotion aware interactive games, neuro marketing, socially intelligents software apps, computers should consider the emotions of their human conversation partners. Speech analytics and 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 speech signals or facial expression signals have disadvantages: using only them is not reliable to detect emotion, especially when people want to conceal their feelings. Compared with facial expression, using physiological signals is a reliable approach to probe the internal cognitive and emotional changes of users. In this study, emotion recognition from Galvanic Skin Response was performed using time domain based features and wavelet approaches Valence and arousal have been categorized and relationship between physiological signals and arousal and valence has been studied using Random Forest machine learning algorithm.

978-1-5090-2386-8/16/$31.00 ©2016 IEEE

II. LITERATURE SEARCH

Emotions regulate the autonomic nervous system, which, in turn, causes variations in the secretion of sweat on the skin's surface, as well as changes in the heart rate and respiration rate [3].

GSR, which is known also as electro dermal activity(EDA) is a low cost, easily captured physiological signal. GSR is a reflection of physiological reactions that generate excitement. Emotional arousal induces a sweat reaction, which is particularly prevalent at the surface of the hands and fingers and the soles of the feet. When people get excited, body sweats, the amount of salt in the skin increases and the skin's electrical resistance also increases. GSR appears sensitive only to the arousal dimension not direction or valence of the emotion involved. Skin conductivity varies with changes in skin moisture level(sweating) and can reveal changes in sympathetic nervous system. Nakasone et al. have used skin conductance and muscle activity for emotion recognition [4]. Nourbakhsh et al. investigated different time and frequency domain features of GSR in multiple difficulty levels of arithmetic and reading experiments [5]. 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 [6].

III. MATERIALS AND METHODS

A. Galvanic Skin Response Signals

In GSR method, the electrical conductance of the skin is measured through one or two sensor(s) usually attached to hand or foot. This resistance decreases due to an increase of perspiration, which usually occurs when one is experiencing emotions such as stress or surprise.

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.

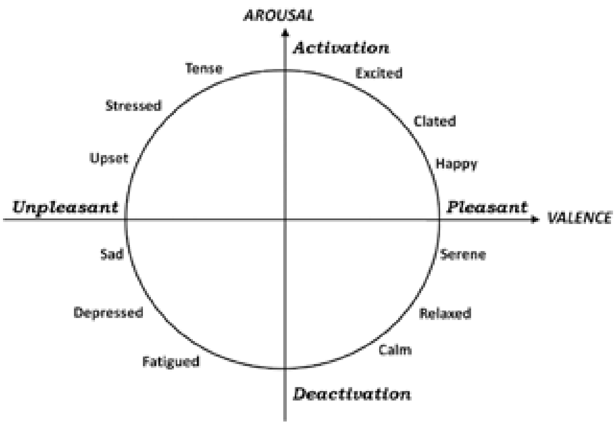

Figure 1. Valence - Arousal Model

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. 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 dataset was first presented by Kolestra et al. [7]. The data was down sampled to 128Hz, EOG artefacts were removed, a bandpass frequency filter from 4.0 - 45.0Hz 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. The dataset contains both EEG and peripheral physiological signals. In this paper, among recorded channels only Galvanic Skin Response(GSR) signals have been considered.

D. FEATURE EXTRACTION

Features from signals have been extracted in the time domain and based on statistics.

GSR signal has been subjected to various length moving windows for feature extraction. In each trial, we have obtained four channels' signals and divide each channel signal into segments (e.g. 20 segments with 3s length per segment). 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, moments, change in signal values have been considered as features. Various attributes have been selected as feature set and relationship between arousal and valence has been studied. Table 1 shows studied feature sets and their attributes. Each feature set has been

Feature Set	Attributes
Feature- 10	*Minimum, Maximum, Average, Standart Deviation, Variance, Skewness, Kurtosis, Median, Zero Crossings, Mean Energy*
Feature - 14	*Feature 10 Set,* *3rd, 4th, 5th, 6th Moment*
Feature - 18	*Featue 14 Set, Mean Absolute Value, Max Scatter Difference, Root Mean Square, Mean Absolute Deviation*
Feature - 22	*Feature 18 Set, 1st Degree Difference, 2nd Degree Difference, 1st Degree Diff Divided with Std Deviation, 2nd Degree Diff Divided with Std Deviation*

Table 1. Feature Sets and Attributes

E. 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). GSR signals taken from 32 subjects all have been used for training and test steps. After feature extraction the signals are classified into classes using Random Forest. Random Forests [8] 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. Random forests achieve high accuracy in a variety of problems, making them versatile choice for many applications. Since only a subset of the features used, random forests capable of handling high dimensional data. Also, a trained model can be used to determine the pairwise proximity between samples. These features make random forests a popular technique in bioinformatics and specialized random forests for these purposes are an active area of research.

Feature Size	Record Size	Class Size	AROUSAL Accuracy No -Conv	AROUSAL Accuracy Convolution	VALENCE Accuracy No-Conv	VALENCE Accuracy Convolution	Window Duration (sn)
10x63	40x32	2	70.78%	70.78%	69.6%	69.6%	1
10x21	40x32	2	71.53%	71.46%	70.54%	71.04%	3
10x12	40x32	2	70.46%	70.76%	69.68%	70.39%	5
10x8	40x32	2	69.6%	70.17%	69.49%	70.23%	8
10x6	40x32	2	69.0%	70.15%	69.32%	69.76%	10
10x5	40x32	2	68.96%	69.45%	69.21%	69.56%	12
10x4	40x32	2	68.75%	68.9%	68.92%	69.07%	15
10x2	40x32	2	68.04%	68.44%	68.21%	68.37%	30
10x1	40x32	2	66.48%	66.48%	65.7%	65.7%	60

Table 2. GSR – Time Domain Statistics Experiments

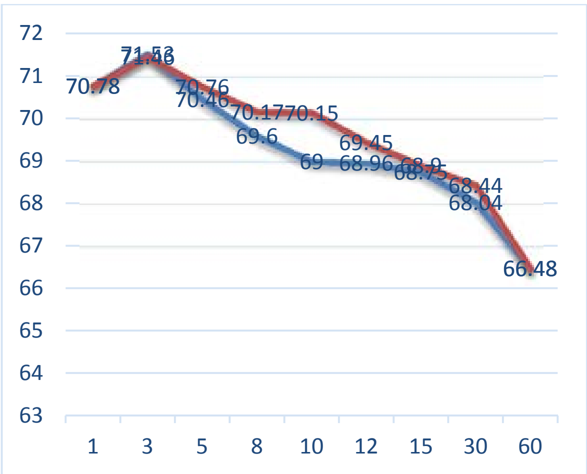

Fig 2. Arousal: Accuracy Rate vs. Window Duration Size

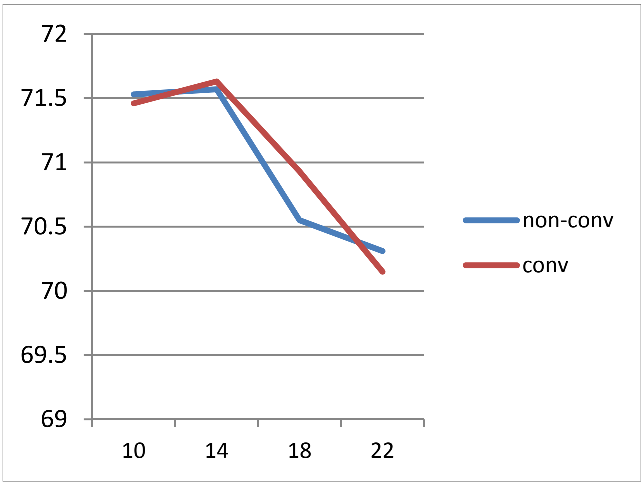

Fig 3: Arousal Accuracy Rate vs Feature Set Size

Class	Wavelet		Time	
Arousal	70.31%	70.39%	71.57%	71.63%
Valence	70.7%	70.31%	71.54%	71.85%
	Non-Convolution	Convolution	Non-Convolution	Convolution

Table 3. GSR – Time Domain Statistics Experiments

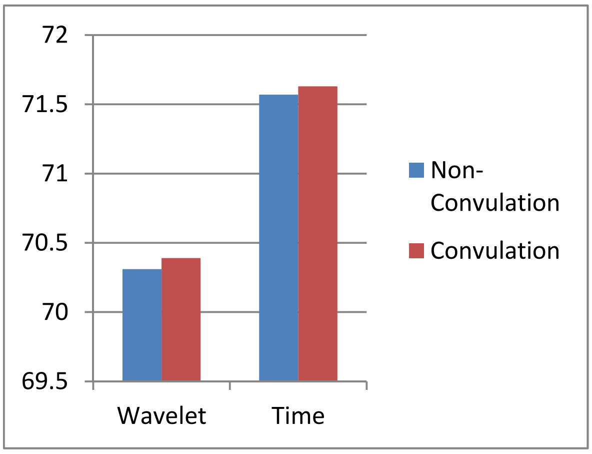

Fig 4: Wavelet and Time based features vs Accuracy Rate

F. EXPERIMENTAL RESULTS AND DISCUSSION

Many approaches have been tested. Tests have been conducted with 10-fold cross validation by using Random Forest machine learning algorithm.

Window Duration Size Tests

Window duration has effect on accuracy rate. Various window size duration between 1 seconds and 60 seconds have been selected. Tests with 3 seconds window duration performed better than other window duration size. Results are depicted in Figure 2.

Feature Set Tests

Feature extraction has effect on accuracy rate. Various feature sets(FS) have been selected. Tests with FS 10, FS 14, FS 18 and FS 22 has been conducted. FS 14 performed better seconds window duration performed better than other feature sets. Results are depicted in Figure 3.

Convolution vs. Non-Convolution Tests

Windows have been slided by collapse or not collapse manner. Overlapped and one second slide duration has performed better compared to non-overlapping window sliding. Figure 2 and Figure 3 confirms that convolution is a better approach to increase accuracy rate.

Wavelet versus Time Based Features Tests

Time based features and wavelet approach has been compared by tests. Time based features performed better as shown in Table 3 and Figure 4.

G. CONCLUSION AND FUTURE WORK

Recognizing arousal and valence values directly from only GSR Signals is a challenge task. We have seen that there is relationship between GSR signals and arousal and valence. In case of we categorize both arousal and valence into two class we have achieved 71.53% and 71.04% accuracy rate for arousal and valence respectively. In the future we are planning to apply data fusion techniques with other physiological signals and apply different machine learning algorithms to increase accuracy rate.

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] Shimmer, "Measuring Emotion: Reactions To Media", Dublin, Ireland, 2015

[4] A.Nakasone, H.Predinger, and M.Ishizuka, "Emotion Recognition from Electromyography and Skin Conductance", in proc. BSI 2005(2005), 219-222

[5] N.Nourbakhsh, Y. Wang, F. Chen, R.. Calvo, "Using Galvanic Skin Response for Cognitive Load Measurement in Arithmetic and Reading Tasks", ACM 2012

[6] 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.

[7] 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.

[8] Ho, Tin Kam. Random Decision Forests (PDF). Procedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–28