2B Circumplex Modeli ile Çok Değişkenli Senkrosıkıştırma Dönüşümü Kullanarak Duygu Sezimi

Emotion Detection Using Multivariate Synchrosqueezing Transform via 2D Circumplex Model

Pinar OZEL[1] and Aydin AKAN[2]
[1] Dept. of Electrical-Electronics Eng.
[2] Dept. of Biomedical Engineering
[1]Nevsehir Hacı Bektas Veli Univ., Nevsehir, Turkey
[2]Izmir Katip Celebi Univ., Izmir, Turkey
pinarozel@nevsehir.edu.tr , aydin.akan@ikc.edu.tr

Bulent YILMAZ[3]
[3] Dept. of Electrical-Electronics Eng.
Abdullah Gul Univ.,
Kayseri, Turkey
bulent.yilmaz@agu.edu.tr

***Özetçe*— Sinyal işleme yöntemlerini kullanarak duygu tespiti ilgi çeken bir alandır. Duygusal modellemede tartışılan bir konu sınıflandırma süreci için kullanılacak optimum öznitelik kümesi elde etmektir. Bu çalışma, EEG sinyallerinin Çok Değişkenli Senkrosıkıştırma Dönüşümü (CDSSD) ile analiz edilmesi yoluyla duygusal durum sınıflandırması için bir yaklaşım önermektedir. CDSSD, çok değişkenli senkrosıkıştırma katsayıları veren bir lokalize zaman-frekans temsilini oluşturmak için bir işlem sonrası tekniğidir. DEAP veri setinden 18 kişi için EEG sinyallerinden bu katsayılar elde edildikten sonra, katılımcıların CDSSD katsayıları ve özdeğerlendirme etiketleri, en yakın komşu, karar ağacı ve topluluk destek vektör makineleri kullanılarak duygusal durum sınıflandırması için kullanılmıştır. Destek vektör makineleri kullanılarak elde edilen doğruluk oranları, yüksek değerlik yüksek uyarılma (YDYU) için % 70.6, düşük değerlikli yüksek uyarılma için % 75.4 (DDYU), yüksek değerlikli düşük uyarılma için% 77.8 (YDDU) ve düşük değerlikli düşük uyarılma (DDDU) için% 77.2'dir**

Anahtar Kelimeler — Duygu Durum Analizi; EEG; Çok Değişkenli Senkrosıkıştırma Dönüşümü

***Abstract*— Emotion detection by utilizing signal processing methods is a challenging area. An open issue in emotional modeling is to obtain an optimum feature set to use for the classification process. This study proposes an approach for emotional state classification by the investigation of EEG signals via multivariate synchrosqueezing transform (MSST). MSST is a post-processing technique to compose a localized time-frequency representation yielding multivariate syncyrosqueezing coefficients. After obtaining these coefficients from EEG signals for 18 subjects from DEAP dataset, coefficients and self-assessment-mannequins (SAM) labels of those subjects are used for emotional state classification by using support vector machines (SVM) nearest neighbor, decision tree, and ensemble methods. The accuracy rate is 70.6% for high valence high arousal (HVHA), 75.4% for low valence high arousal (LVHA), 77.8% for high valence low arousal (HVLA), and 77.2% for low valence low arousal (LVLA) cases using SVM.**

Keywords— EEG, Emotion Recognition, Multivariate Syncyrosqueezing Transform.

I. INTRODUCTION

Emotions play an important role in our daily lives because they affect social interaction, cognition and behavior. When we consider emotions in terms of affective computing in which human-computer interaction (HCI) is coordinated, it is clear that the information belonging to emotions shows significant improvement. HCI needs to have emotional intelligence like human-human interaction. As such, emotional analysis methods are investigated using physiological signals such as face expression [1], speech [2] and heart rate [3] and galvanic skin response (GSR) [4] in addition to central nervous system signals. However, electroencephalography (EEG), magnetoencephalogram (MEG), position emission tomography (PET) and functional magnetic resonance imaging (fMRI) that are categorized as central nervous system signals provide more reliable results compared to the physiological signals obtained from the autonomic nervous system (ANS) [5].

Electroencephalography (EEG) is a basic method for realizing a brain computer interface. The brain signals spatially collected through appropriate EEG sensors are processed using time and frequency domain methods. Normally, the EEG data are not stationary by nature and therefore it is more appropriate to apply time-frequency analysis techniques. One of the first conceivable methods is the Short-Time Fourier Transform (STFT), but STFT is not very determinative because it is subject to Heisenberg uncertainty and the resolution of the time-frequency representation is limited by the choice of the window

978-1-5386-6852-8/18/$31.00 ©2018 IEEE

and cannot guarantee the stability of neural data in each short-time period [6][7]. On the other hand, wavelet-based transforms are assumed to be data adaptive in signal processing methods. However, the disadvantage is that it uses a basic function known as the main wavelet for signal decomposition and faces the problem of time-frequency resolution, that is, lower frequency resolution at higher frequencies and higher resolution at lower frequencies may be the issue. Additionally, wavelet analysis also depends on the choice of the main wavelet. Determination of the desired main wavelet without being coordinated with the analyzed signal can lead to erroneous results [8]. Then, recent advances in time-frequency algorithms have provoked the development of high-resolution time-frequency methods such as the empirical mode decomposition (EMD) and the synchrosqueezing transform (SST) [7]. EMD is a strategy for examining signal into its intrinsic mode functions. On the other hand, synchrosqueezing transform, which is an EMD like approach, is a mixture of wavelet and reassignment technique extensively. The goal of SST is to reassign the energies of STFT and continuous wavelet transform (CWT) or similar time-frequency algorithms, such that the subsequent energies of coefficients are intense around the instantaneous frequency curves of the modulated oscillations [8].

Emotion detection utilizing the multichannel electroencephalography signals (EEG) is a relatively new methodology to assess the human-computer interaction. In psychology, emotions can be expressed by continuous labels such as the valence and arousal plane proposed by Russell's circumplex model [9]. In this plane, while arousal covers the area between ineffective (e.g. uncontrolled, squeezed) to effective (e.g.: careful, enthusiastic), valence covers unpleasant plane (e.g., sad, stressed) to pleasant (e.g., cheerful, amused) as shown in Figure 1. Though it is a fact that valence and arousal co-vary in the reality [10], late examinations propose that these two dimensions collaborate and should be considered when contemplating emotional modeling in biomedical signal processing [11] [12].

Frantzidis et al. [13] presented a methodology for the electrophysiological data into four emotional situations gathered while the volunteers in the experiment were seeing emotional pictures passively chosen from IAPS. It receives the independence of two emotive dimensions named as arousal and valence. Four emotional states were arranged as high valence high arousal "HVHA", low valence high arousal "LVHA", high valence low arousal "HVLA", and low valence low arousal "LVLA" using the Mahalanois distance (MD) based classifier and support vector machines (SVMs). Petrantonakis et al. [14] applied a multi-dimensional directed information analysis among distinct EEG sites from the two opposite brain hemispheres and conducted a comprehensive classification process by utilizing within VA space as four affective statements. Liu et al suggested a new fractal-based method using for emotion detection from EEG operating VA space [15]. Mert and Akan studied both high/low emotional dimensions and four emotional states using Multivariate Empirical Mode Decomposition for feature extraction method by modelling via valence and arousal dimensions [16]. Lin et al. proposed two types of multi-class SVM algorithms for emotion recognition with organization of VA space after extracting features of EEG signals via Short-Time Fourier Transform [17].

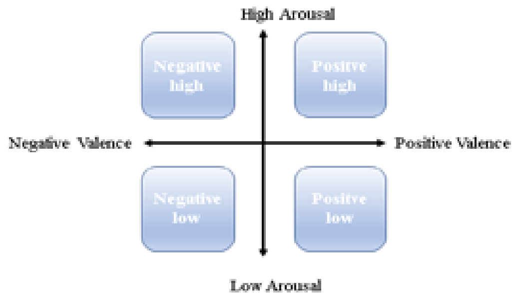

Figure. 1. Two-dimensional model by arousal and valence for emotion recognition.

The goal of the present study is to perform a 2D valence-arousal model for the classification of emotional states via analysis of EEG signals and to introduce new alternatives in the classification phase to increase the classification performance according to the new methods used in time frequency domain. Multivariate synchrosqueezing transform (MSST) is preferred to utilize the knowledge of inter-channel dependencies that may arise from multichannel data. Additionally, the synchrosqueezing coefficients obtained at the end of MSST algorithm are proposed to present EEG signals multivariate inherent for input feature vector in the classification step. Section II describes multivariate synchrosqueezing transform concepts, dataset and our proposed approach. Section III presents the results obtained by applying the proposed strategies and discusses the conclusions of this study.

II. METHODOLOGY

A. Multivariate Synchrosqueezing Transform

Synchrosqueezing transform is a method designed to extract and compare the oscillation components of the signals in complex systems and to obtain clearer indications on the time-frequency axis. It is a post-processing technique applied on continuous wavelet transform to produce localized time-frequency representations of non-stationary signals. The general form of the signal x (t) is specified as follows:

$$y(t) = \sum_{n=1}^{N} y\,(t) + c(t) \qquad (1)$$

Accordingly, each component $y_n(t) = B_n(t) \cos(\emptyset_n(t))$ is an oscillation function to represent the noise, c (t), having the amplitude and frequency varying with time. The aim, here, is to obtain the amplitude factor $B_n(t)$ and the instantaneous frequency (IF) $(\emptyset_n(t))$ for each k.

Synchronization occurs in three steps. First, the continuous wavelet transforms $X_y(a, b)$ of y (t) is calculated. Continuous wavelet transform is a predictive-based algorithm that identifies the corresponding oscillation components through time-frequency filters known as wavelets. The wavelet ψ (t) is the final oscillation function that is convolved with the signal y (t). Wavelet coefficients X_y (a, b) are given by,

$$X_y(a,b) = \int a^{-1/2}\psi\left(\frac{t-b}{a}\right)y(t)dt. \qquad (2)$$

In this way the wavelet coefficients can be seen as the outputs of a bandpass filter. The scale factor 'a' scrolls the bandpass filter in the frequency domain and changes the bandwidth of the bandpass filter. In the second step, for each scale time pair (a, b), the instantaneous frequency x_y(a, b) can be estimated as given below.

$$x_y(a,b) = -iX(a,b)^{-1}\frac{\partial X(a,b)}{\partial b} \qquad (3)$$

Finally, this estimate is used to squeeze S_y by the reassignment method. As a result, the synchrosqueezing representation S_y (x, b) is obtained [20,21,22]. That is, the wavelet coefficients obtained contain the same instantaneous frequencies and are then combined by the synchrosqueezing (SST) method. For a set of wavelet coefficients X (a, b).

$$S(x_l,b) = \sum_{a_n:|x_y(a_n,b)-x_l|\le\Delta x/2} X(a_n,b)a^{-3/2} \vartriangle a_n \qquad (4)$$

Multivariate SST Algorithm (MSST):
For the given M-channel multivariable signal y(t), SST is applied to according to the channel to obtain S_m (x, b) coefficients. For the time-frequency representation, a series of sections along the frequency axis are determined, the instantaneous frequency and the instantaneous amplitude for each frequency section are calculated. In the next step, Multivariate instantaneous frequency and instantaneous amplitude are calculated, respectively. As a result, the multivariate synchrosqueezing coefficient $S^{multi}(x, b)$ is determined [19].

B. Dataset

DEAP [20] is a database that is approved to be utilized for academic research, and it consists of EEG signals from various participants who were exposed to emotional videos. The participants watched music recordings and assessed each video as for arousal, valence, like/dislike, dominance, and familiarity completing self-assessment mannequins (SAM) and score it with the numbers 1– 9 for emotional states after each video is viewed immediately. The database is included multichannel EEG signals from 32 participants (16 men and 16 women, whose ages go from 19 to 37, mean: 26.9). The signals utilized as a part of this research were a subset of DEAP dataset. Only 40 one-minute long music recordings with different content were chosen. The EEG signals were recorded from Fp1, AF3, F3, F7, FC5, FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, C4, T8, CP6, CP2, P4, P8, PO4, and O2 electrodes.

III. The Proposed Method

The EEG signals recorded as the DEAP data set was used in this study. EEG signals of 18 subjects and 40 video recordings were considered. Each video-clip duration was 4032 samples. The dataset was given as the input to the MSST algorithm consisting of EEG channels as the right frontal- weighted lobe (Fp2, AF4, F4, F8, FC2, FC6, T8, C4), the left frontal-weighted lobe (Fp1, AF3, F3, F7, FC1, FC5, T7, C3), the right and left frontal-weighted lobe differences and 2 central channels (Fz, Cz) for each subject. Because of the algorithm, multivariate synchrosqueezing coefficients were obtained and a 26x2017x4032 matrix was obtained for each video recording. When these were given as the input to the singular value decomposition (SVD) algorithm for dimension reduction, a one-dimensional matrix was obtained for each video whose dimension was 52416x1. Since the array obtained here contained complex numbers, an absolute number was formed by taking an absolute value. For the data of the 18 people used in the study, an input data of 720 * 52416 was created (each video for the data of the 18 subjects used is arranged in the lower sub-array and the input matrix was formed.

For each label (valence-arousal) correspondingly, it was determined that the values varying from 1 to 9 have high or low values according to the 5th scale <5, and the last label values were set to 0 or 1. 72 x1 with two output data according to their arousal and valence labels were exposed to logical mathematics. So, if both valence and arousal labels were equal to 1, it was set as high valence high arousal "HVHA". If valence was equal to 0 and arousal was equal to 1, it was set as low valence high arousal "LVHA" and similarly, high valence low arousal "HVLA", low valence low arousal "LVLA" were determined. With the help of this logical computation, four emotional states were obtained as HVHA, LVHA, HVLA, LVLA and set as four label outputs. Finally, classification process is executed after the suitable features are specified and the feature vector is created. The classifiers are trained at first and then determine the class of an unknown feature vector. To determine the performance of the classifiers, the training and test data sets may be determined using cross validation procedures. Hence, some of the data is used for the training phase, and the rest is used for testing. The performances of the selected classifiers in MATLAB Classification Learner Tool package were evaluated and reported. The classification approaches used in this study are support vector machines, nearest neighbor, decision tree, and ensemble methods.

IV. Results and Discussion

In this study, the DEAP dataset for emotional state analysis was used to classify four emotion states as VA space as HVHA, HVLA, LVLA, LVHA. In the studies that used EEG signals, the effects on the classification performance of stimuli were compared via classical time, frequency, time-frequency features. In this study, as the input feature vector multivariate synchrosqueezing coefficients were used and valence and activation labels were combined via utilizing "and operator" in logical mathematics. In Table I, the accuracies for each class of emotional states were given. The accuracies were over 70% for most the classification methods. We note that the success levels were satisfactory. We believe that we can get higher accuracy rates by selecting better dimension reduction or classification approaches or other algorithmic changes. As a limitation, in this study we did not take the advantage of dominance level contribution.

In the future, we have the idea of modeling emotions via valence-arousal-dominance space utilizing multivariate synchrosqueezing coefficients. Our study is also important in terms of focusing on valence or arousal labels one by one. For example, in neuromarketing studies, it is significant to perceive the contribution of valence label due to its positive-negative scale. Additionally, for arousal labels applications like movie examinations, it is necessary to see the contribution of arousal label in terms of how much an emotion (i.e. horror) is increased or decreased. However, we realize that it is noteworthy to combine all these spaces by overlapping different emotions in one-dimensional emotional space (high-low valence and high-low arousal separately) or 2D dimensional emotional space (VA space). Our results imply that just the emotional area as HAHV, HALV, LALV, LAHV in 2D emotional space is not the different emotions such as sad or happiness directly and separately. We propose this methodology in VA space for more basic applications in need of four distinct emotional states. Our future studies will include approaches to determine emotions in VAD space.

Table 1. Accuracy rates for four emotional states.

	HAHV	HALV	LALV	LAHV
Support Vector Machine Linear SVM	70.6%	77.8%	77.2%	75.6%
Nearest Neigbor Medium k-NN	70.7%	77.6%	76.3%	75.4%
Decision Tree Medium Tree	64.4%	74.6%	72.9%	74.0%
Ensemble Bagged	65.6%	75.8%	75.8%	73.6%

The results obtained with the proposed method are promising. In the literature related to DEAP database the results were 62% accuracy for high/low activation discrimination, 57% for high/low valence discrimination. For 2D VA space applications, the accuracy performances decreased. Our one-dimensional accuracies were also better than the DEAP database.

REFERENCES

[1] Gunes , H. , Piccardi , M. , "Fusing face and body display for bi-modal emotion recogniiton: Single frame analysis and multi-frame post integration". ACII'05 Proceedings of the First international conference on Affective Computing and Intelligent Interaction, Beijing, 2005.

[2] Ayadi, M.E., Kamel, M.S., Karray, F., "Survey on speech emotion recognition: Features, classification schemes, and databases". Pattern Recognition, pp. 572–587, 2011.

[3] Yu, S.N., Chen, S.F., "Emotion state identification based on heart rate variability and genetic algorithm".37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy , 2015.

[4] Liu, M. , Fan, D., Zhang, X., Gong, X., "Human Emotion Recognition Based on Galvanic Skin Response Signal Feature Selection and SVM " . International Conference on Smart City and Systems Engineering (ICSCSE), Hunan, China, 2016.

[5] Petrantonakis , P. C. , Hadjileontiadis , L. J. , "Emotion recognition from Brain Signals Using Hybrid Adaptive Filtering and Higher Order Crossings ". IEEE Tran.on Affective Computing , vol.2, pp. 81-97, 2010.

[6] Mallat, S., A Wavelet Tour of Signal Processing, Newyork: Academic Press, 1998.

[7] Auger, F. , Flandrin, P. , Lin, Y. T. , McLaughin, S. , Meignen, S. , Oberlin, T. ,Wu, H. T. , "Time Frequency Reassignment and Synchrosqueezing ". IEEE Signal Processing Magazine, pp.32-41,2013

[8] Daubechies, I., Lu,J., Wu,H.T., "Synchrosqueezed wavelet transforms: an empirical mode decompositionlike tool". Applied and Computational Harmonic Analysis, pp. 243-261, 2011

[9] Russell, J. A., "A Circumplex Model of Affect" Journal of Personal and Social Pschology, vol. 39, pp. 1161-1178, 1980.

[10] Lang, P.J., Bradley, M.M., Cuthbert, B.N.: Motivated attention: affect, activation, and action. In: Lang PJ, Simons RF, Balaban M (eds), " Attention and orienting: sensory and motivational processes. Erlbaum Associates," Erlbaum Associates, 1997.

[11] Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J. "Brain potentials in affective picture processing: covariation with autonomic arousal and affective report". Biol Psychol, vol. 52, p. 95–111, 2000.

[12] Robinson, M.R., Storbeck, J., Meier, B.P., Kirkeby, B.S. "Watch out! That could be dangerous: valence-arousal interactions in evaluative processing". Pers Soc Psychol Bull, vol. 30, p. 1472–1484, 2004.

[13] Frantzidis, C. A., Bratsas, C., Papadelis, C. L., Konstantinidis, E., Pappas, C., Bamidis, P.D. "Toward emotion aware computing: An integrated approach using multichannel neurophysiological recordings and affective visual stimuli,". IEEE Transactions on Information Technology in Biomedicine, vol. 14, pp. 589-597, 2010.

[14] Petrantonakis, P.C., Handjileontiadis, L.J. "A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG based emotion recognition". IEEE Trans Inf Technol Biomed. 2011 Sep;15(5):737-46

[15] Liu, Y., Sourina, O., Nyuyen, M.K. "Real Time EEG-Based Emotion Recognition and its Applications". Transactions on Computational Science XII, 20011.

[16] Mert,A.,.Akan,A., "Emotion recognition from EEG signals by using multivariate empirical mode decomposition". Pattern Analysis and Applications, pp. 1-9, 2016.

[17] Lin,Y.P., Wang, C.H, Wu, T.L. , Jeng , S.K. Chen., , J.H. "EEG-Based Emotion recognition in musiz listening : a comparision of schemes for multiclass support vector mmachine". Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on

[18] Ahrabian, A., Looney, D., Stanković, L., Mandic, D.P.: "Synchrosqueezing-based time frequency analysis of multivariate data.". Signal Processing, pp. 331-341, 2015.

[19] Daubechies, I., Lu,J., Wu,H.T., "Synchrosqueezed wavelet transforms: an empirical mode decompositionlike tool". Applied and Computational Harmonic Analysis, pp. 243-261, 2011

[20] Koelstra, S., Mühl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., Patras, I., "DEAP: A database for emotion analysis; using physiological signals". IEEE transactions on affective computing,, vol 3, pp. 18-31, 2012.