Application of Multiscale Entropy on EEG signals for emotion detection.

Kostas Michalopoulos and Nikolaos Bourbakis

Abstract— **Multiscale Entropy (MSE) is a method that measures the temporal regularity of time series across different time scales. In this paper, we study the use of Multiscale Entropy for the detection of different emotion states. We apply MSE in Electroencephalographic (EEG) recordings of subjects watching musical videos, selected to elicit specific emotions. MSE is able to uncover significant differences in the temporal organization of the EEG during events that elicit emotions with low/high valence and arousal.**

I. Introduction

Emotion detection or affect detection from physiological signals is a challenging problem which has attracted a lot of attention [1]–[3]. The goal is to detect and identify the emotional or affective state of the subject using inputs from different modalities. The main challenge is that emotion is subjective construct that cannot be directly observed and measured. We have to rely on the subjective evaluation of the individual regarding the affective influence of an object or situation.

For this reason different scales for categorization of emotions have been proposed to facilitate the reporting of the elicited emotions. In this study the valence-arousal scale has been used for rating [4], [5]. In this model the emotional states are characterized by two dimensions, valence and arousal and they can be mapped to a plane with arousal as the horizontal, and valence as the vertical axes. Arousal maps emotions ranging from inactive to active while valence ranges from unpleasant to pleasant. The emotion labels of this scale can be seen in Figure 1.

EEG measures the electrical activity on the scalp of the head reflecting the underlying brain activity. There are different studies that used EEG to classify the emotional states during an affective experiment. The extracted features are mainly the spectral power of the different bands and channels [5], [6]. Such features, capture only the linear interactions and are not capable to capture the dynamical properties and of the brain and the nonlinear interactions between the activated sources.

Methods that quantify the temporal structure and complexity of the EEG signal have provide useful insights in the study of heart rate variability and more recently neurophysiological data [7], [8]. Multiscale entropy is a measure of the temporal regularity of a signal, evaluated at different time scales [7]. The evaluation of regularity in larger time scales allow the detection of long-range correlations that can provide useful insights regarding the mechanisms underlying physiological dynamics. Indeed, MSE has been

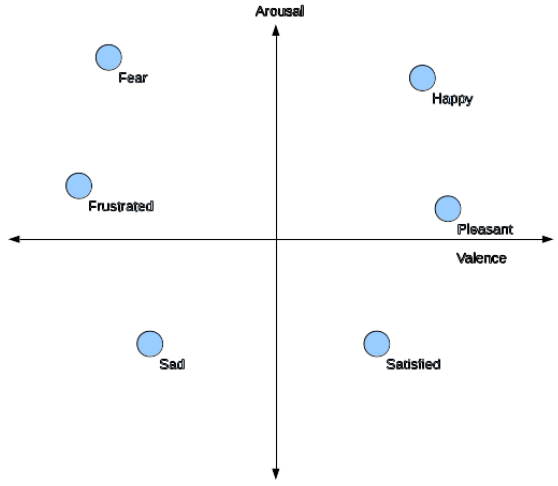

Figure 1: Illustration of the valence/arousal model for emotion categorization, adapted from [4].

successfully applied in the nonlinear analysis of EEG and fMRI signals for studying the effects of aging Alzheimer's disease, schizophrenia and autism [8]–[10].

In this paper we study changes in the complexity temporal structure of the EEG during the presentation of musical videos used to elicit different emotions. The underlying assumption is that different brain mechanisms are involved during an emotional state and their interactions will change the temporal patterns of the recorded EEG.

II. Materials and Methods

A. Multiscale Entropy

Multiscale entropy is a measure of time series complexity that has used for the study of biological signals and recently has been applied in the study of neuroimaging signals, like fMRI and EEG [7], [11]. It is a measure of regularity of the time series, measured by Sample Entropy (SE), across time scales. The coarse grain procedure is used to get the signal for different time scales.

B. Sample Entropy

Sample entropy was introduced by [12] as an alternative to approximate entropy that is less sensitive to the length of the time series and the more consistent for a range of its input parameters. SE measures the similarity or predictability of the signal over time. Higher values of SE indicate lower predictability or higher complexity.

Kostas Michalopoulos is with Wright State University, Center of Assistive Research Technologies(michalopoulos.2@gmail.com)

Nikolaos Bourbakis is with Wright State University, Center of Assistive Research Technologies (nikolaos.bourbakis@wright.edu).

978-1-5090-4179-4/17/$31.00 ©2017 IEEE

The SE of time series $x = \{x_1, x_2, \dots, x_N\}$ of length N is calculated using the following procedure. The signal is divided in vectors of length m $u_m(i) = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}, 1 \le i \le N - m + 1$ and the number of vectors $B_i^m(r)$ that satisfy $d(u^m(i), u^m(j)) \le r$ are counted, excluding self-comparisons. The distance d is the maximum distance between the samples of the two vectors and is defined as follows:

$$d_{ij} = \max\{|(u_i^m(k) - u_j^m(k))|\}: 0 \le k \le m - 1 \quad (1)$$

For vectors of length m, the quantity $B^m(r) = (N - m)^{-1} \sum_{i=1}^{N-m} B_i^m(r)$ is defined. The same quantity is calculated for the vectors of size m+1 defined as $A^m(r) = (N - m - 1)^{-1} \sum_{i=1}^{N-m} B_i^{m+1}(r)$. Finally, the Sample Entropy of the time series is defined as:

$$SE(m, r, N) = -\ln\left(\frac{A^m(r)}{B^m(r)}\right) \quad (2)$$

SE is therefore, the conditional probability that two similar vectors of length m from the data series *x* will remain the same (within the tolerance *r*) when the next point is also included in the comparison.

C. Coarse-Graining procedure

To evaluate the SE under different time scales, a coarse graining procedure is applied to the original signal and the SE is evaluated for each derived waveform. The signal is divided in non-overlapping windows of length τ, where τ is the current scale. A new time series is generated by taking the average within each window as follows:

$$y_j^t = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, 1 \le j \le N/\tau \quad (3)$$

For scale τ=1, we have the original time series. The SE is calculated for each coarse signal and plotted against the time scales.

D. Data Description

We used the public available data from the DEAP database. We used the preprocessed dataset which is down sampled to 128Hz. The dataset contains the physiological recordings of Photoplethysmography, Electromyography, EEG, and Galvanic skin impedance [5]. The experimental procedure involved an initial 2 minute baseline recording after which 40 videos from a pool of 120 were presented to the subject. Each trial consisted from a 2 second preparatory period, followed by a 5 second baseline recording and finally the display of the music video lasting 1 minute. At the end of each trial the subjects performed a self-assessment task by providing the levels of valence, arousal, liking and dominance. In this study we took under consideration only the valence-arousal ratings. The self-assessment levels for valence and arousal were ranging from 1 to 9 [5].

EEG was recorded at 512Hz using 32 electrodes arranged according to the 10-20 system. Thirty two subjects participated in the experiment, 50% of them were male [5]. The EEG data were common average referenced and down sampled to 256Hz. Eye artifacts were removed using Blind Source Separation and the last 30 seconds of each trial were kept for analysis. The final preprocessed dataset was further down sampled to 128Hz. In this study the final pre-processed dataset was used.

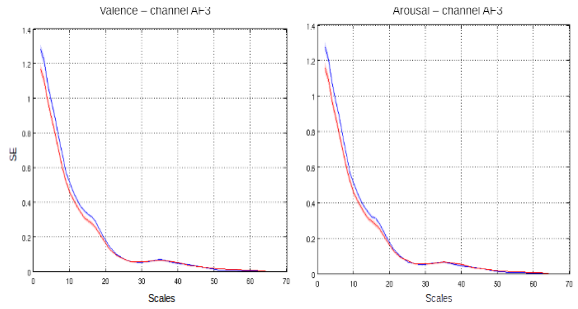

Figure 2: Example of the MSE from a single subject for channel AF3.The blue line is the mean MSE for high valence/arousal trials while the read one is the mean MSE for low valence/arousal trials.

III. Results

We divided the trials of each subject in two groups based on their level of valence and arousal. Using this scheme we performed the analysis and compared trials with low versus high valence, and low vs high arousal. The main objective is to relate changes in the complexity of EEG to low/high levels of either category. Based on previous results on the dataset in [5], the EEG activity correlated to valence is manifested differently from the activity correlated to arousal. The topographic distribution is also different with different channels presenting significant activations between the two conditions. Thus, for the study of MSE we will consider the two cases separately.

MSE was calculated independently for each subject. Each channel was normalized to zero mean and unit variance before the calculation of MSE. The normalization procedure was performed across all trials. Then, for each trial (40 trials) and channel (32 channels) the MSE was calculated for 28 scales, covering the frequencies 1-32Hz. In accordance with other studies we used m=2 and r =0.25 [7]. The relation between the time scale τ and frequency is calculated by the following equation:

$$f_\tau = \frac{f_s}{(2*\tau)} \quad (4)$$

We divided the scales based on their corresponding frequency to bands, using equation 4. We used thirteen scales from those corresponding to the delta band [0.5-4Hz], 7 scales from the theta band [4-7Hz], 5 scales from the alpha band [8-13Hz], and 3 from the beta band [16-31Hz].

We grouped the trials of each subject based on their valence/arousal levels. We used the middle of the 9 level scale as the threshold for categorizing a trial as low/high arousal/valence. For each channel, a Receiver operator characteristic (ROC) analysis was performed on each scale, by comparing the SE levels between trials of low and high arousal/valence. For each scale the probability that the area under the curve (AROC) is significantly different from chance (0.5) was determined using the Mann-Whitney U test [13]. In Figure 2, the MSE from a single subject from channel AF3 is presented.

The mean AROC for the four bands was calculated across subjects for each channel. The topography of the AROC values

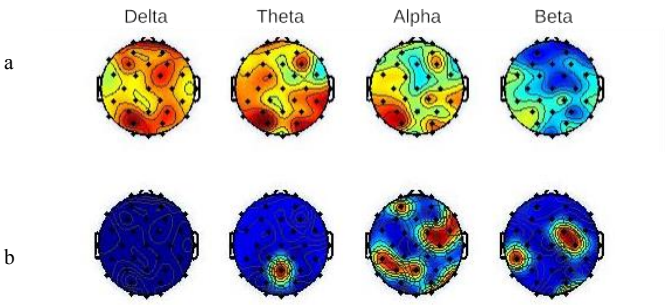

Figure 3: The topography of the mean AROC (top) and the channels with significant activity (p<0.01, bottom) for valence.

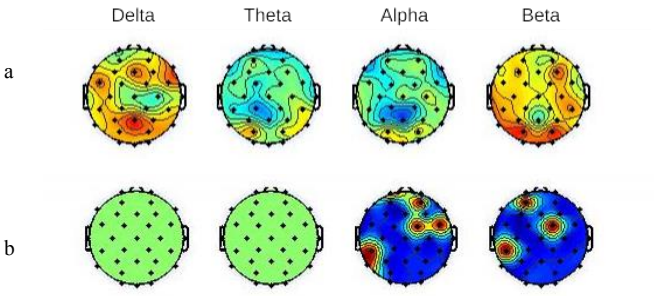

Figure 4: The topography of the mean AROC (top) and the channels with significant activity (p<0.01, bottom) for arousal.

for the valence and arousal can be seen in Figures 3a and 4a, respectively. To determine the significant channels, we combined the individual p-values from the Mann-Whitney test using the Fisher's method. The significant channels for each case can be seen in figures 3b and 4b. We chose a significance level of $p<0.01$ for the channels marked in the figures.

Tables I and II present the significant channels ($p<0.01$) for the low/high valence and the corresponding AROC values. Since the Mann-Whitney U statistic only indicates that the AROC is significantly different from 0.5 the max and min AROC values are presented for the given channel for each band. We can see that the mean AROC values are close to 0.5, even for significant channels. This result is in line with the spectral power analysis in [5], where the mean correlation of the EEG bands to the emotion ratings is low while for the individual subjects high positive or negative correlations were observed. The variability between subjects is high and as a result higher arousal/valence is expressed differently between subjects leading to either increase or decrease of MSE given the emotional state. This is apparent in tables I and II where the high AROC and low AROC values indicate very good discriminative power individually.

Alpha and beta band are the most discriminative bands in both cases. There are channels that exhibit significant changes in both cases. For the alpha band, it appears that the same frontal region is significantly modulated by both valence and arousal. It appears that more channels present changes in their entropy for the valence case than the arousal, concentrated in the posterior electrodes.

Beta band, presents significant activity in fewer electrodes than the alpha band and the activity appears to be concentrated more frontally for the case of arousal. For the other two bands only the theta band contains one active electrode for the case of valence. For the arousal, delta and theta bands do not present significant activity.

The subject variability is high and the low mean AROC values verify this observation. Although, individually we have scales with very high (low) AROC values it appears that the change in the temporal structure of the EEG depends on the individual subject. Thus, we have subjects that the EEG complexity increases with high valence events and subjects that present increases in complexity with low valence events.

TABLE I. CHANNELS WITH SIGNIFICANT AROC FOR VALENCE

Alpha				**Beta**			
	Mean AROC	***Min AROC***	***Max AROC***		***Mean AROC***	***Min AROC***	***Max AROC***
AF3	0.48	0.14	0.78	CP5	0.48	0.22	0.75
CP5	0.51	0.20	0.80	FC2	0.4	0.20	0.73
P3	0.51	0.19	0.76	C4	0.46	0.22	0.75
PZ	0.47	0.18	0.78	O2	0.48	0.24	0.80
F8	0.47	0.19	0.72				
FC6	0.47	0.12	0.77				
FC2	0.47	0.18	0.79				
C4	0.48	0.22	0.76				
P8	0.50	0.17	0.80				
O2	0.51	0.21	0.80				

TABLE II. CHANNELS WITH SIGNIFICANT AROC FOR AROUSAL

Alpha				**Beta**			
	Mean AROC	***Min AROC***	***Max AROC***		***Mean AROC***	***Min AROC***	***Max AROC***
FP1	0.47	0.21	0.84	AF3	0.51	0.18	0.84
CP5	0.48	0.06	0.86	CP5	0.49	0.18	0.83
P7	0.49	0.19	0.77	FC2	0.49	0.17	0.75
AF4	0.49	0.21	0.76				
FC6	0.48	0.16	0.80				
FC2	0.48	0.17	0.78				

IV. DISCUSSION

In this study we examined the use of MSE for the detection of emotional states. MSE reveals changes in the temporal structure of the EEG modulated by the elicited emotion. The ROC values indicate that MSE can be used for the discrimination of emotional states, although the high variability between subjects would require the generation of individual classifiers. For future work, it is interesting to evaluate the effect of the MSE parameters, namely the length of the vector m and the tolerance r. In this study the effects were studied for each channel separately. Using multivariate

techniques for the evaluation of MSE in multiple channels would possibly provide better separation results per subject.

The subject to subject variability observed in this study is in line with previous studies on the same dataset [5]. The use of MSE provides insight to the mechanisms that take place during the different emotional states. Using MSE we observe that certain subjects present a significant increase in the temporal complexity during high valence/arousal events, while in others this effect is reversed This could be suggestive of common mechanism modulating the temporal interactions of the EEG signal and which is selective for certain emotions depending on the subject.

REFERENCES

[1] M. Mauri, V. Magagnin, P. Cipresso, L. Mainardi, E. N. Brown, S. Cerutti, M. Villamira, and R. Barbieri, "Psychophysiological signals associated with affective states.," Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf., vol. 2010, pp. 3563–6, 2010.

[2] A. Sloman, "Review of Affective Computing," AI Mag., vol. 20, no. 1, pp. 127–133, 1999.

[3] C. A. Frantzidis, C. Bratsas, M. A. Klados, E. Konstantinidis, C. D. Lithari, A. B. Vivas, C. L. Papadelis, E. Kaldoudi, C. Pappas, P. D. Bamidis, C. A. Frantzidis, C. Bratsas, M. A. Klados, E. Konstantinidis, C. D. Lithari, C. Pappas, and P. D. Bamidis, "On the Classification of Emotional Biosignals Evoked While Viewing Affective Pictures : An Integrated Data - Mining - Based Approach for Healthcare Applications," IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 2, 2010.

[4] J. POSNER, J. A. RUSSELL, and B. S. PETERSON, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," Dev. Psychopathol., vol. 17, no. 3, pp. 715–734, Sep. 2005.

[5] S. Koelstra, C. Mühl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; Using physiological signals," IEEE Trans. Affect. Comput., vol. 3, no. 1, pp. 18–31, 2012.

[6] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," IEEE Trans. Affect. Comput., vol. 5, no. 3, pp. 327–339, 2014.

[7] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys., vol. 71, no. 2, 2005.

[8] M. O. Sokunbi, "Sample entropy reveals high discriminative power between young and elderly adults in short fMRI data sets," Front. Neuroinform., vol. 8, no. July, pp. 1–12, 2014.

[9] A. Catarino, O. Churches, S. Baron-Cohen, A. Andrade, and H. Ring, "Atypical EEG complexity in autism spectrum conditions: A multiscale entropy analysis," 2011.

[10] V. Jelic and A. Nordberg, "Early diagnosis of Alzheimer disease with positron emission tomography," Alzheimer Dis. Assoc. Disord., vol. 14, no. SUPPL. 1, pp. S109–S113, 2000.

[11] M. U. Ahmed, L. Li, J. Cao, and D. P. Mandic, "Multivariate multiscale entropy for brain consciousness analysis," in Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2011.

[12] J. S. Richman, J. R. Moorman, A. J. Physiol, and H. Circ, "Physiological time-series analysis using approximate entropy and sample entropy," Am. J. Physiol. - Hear. Circ. Physiol., vol. 278, no. 6, pp. 2039–2049, 2008.

[13] H. B. Mann and D. R. Whitney, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," Ann. Math. Stat., vol. 18, no. 1, pp. 50–60, Mar. 1947.