Proceedings of the 11th International Conference on Applied Informatics
Eger, Hungary, January 29–31, 2020, published at http://ceur-ws.org

Music Stimuli Recognition from EEG Signals with EMOTIV EPOC Headset

Jozsef Suto, Zsanett Jambor

University of Debrecen, Department of IT Systems and Networks, Debrecen, Hungary
suto.jozsef@inf.unideb.hu, jamborzsanci@gmail.com

Abstract

It is well-known that, music stimuli have powerful emotion trigger effect. When people listening to music, music induces motor system activities in their brain. Therefore, music can be used as a potential stimulus in electroencephalogram (EEG) based emotion research. The goal of emotion recognition is to explore how different kinds of stimuli (e.g. music) from the world around us influence our brain waves. In previous works, the determination of emotional states has based on subjects' feedback. However, this approach is unreliable almost in all cases because emotional states are changing rather slowly and they are equivocal. In this study we try to recognize music-induced electroencephalogram patterns by shallow artificial neural network from the popular EMOTIV EPOC+ sensor's signals. This article presents the data acquisition conditions; the efficiency of the neural network with different hyper-parameters; and the effectiveness of EMOTIV EPOC+ over the Neurosky Mindwave device.

Keywords: Artificial neural network, digital filtering, EMOTIV EPOC, feature engineering, music stimuli

MSC: 68T10, 92C55

1. Introduction

In emotion analysis, there are three different approaches. The first type focuses on the analysis of facial expressions or speech. In the second approach periphery physiological signals such as electrocardiogram, skin conductance respiration, and pulse are used for emotion prediction. The third approach is based on brain signals such as EEG or functional magnetic resonance imaging (fMRI) [11]. Among those possibilities, EEG signals have become widely used as data source in the emotion and mental state recognition research. After the appearance of some commercial EEG reader devices, a significant part of EEG related papers dealt with this topic [13, 5].

More different stimuli effects can be used to elicit emotions. Among them the most popular are images, sounds, and videos. In the study of Yuvaray et al. [12] the authors applied visual and audio stimuli to elicit participants' emotional responses. They used two classifier algorithms (k-nearest neighbor and support vector machine) with high order spectra (HOS) features to compare the emotional states of healthy and Parkinson's disease subjects. In their experiment, at the end of all EEG recordings, participants had to fill a questionnaire to state the status of emotions they felt. In the questionnaire 6 elementary emotions (disgust, fear, sadness, etc.) could be selected and the strength of feeling also had to be marked on a 5-point scale. Meza-Kubo et al. [3] performed user experiment evaluation with the EMOTIV EPOC+ headset. In this experiment, elder couples played a simple computer game and their emotions (natural, pleasant, unpleasant) were predicted in three different ways (self-reported, qualitative analysis and EEG signal classification). Their result also illustrated the issue of self-reported emotion state determination. Lin et al. [2] used music pieces from the Oscar's film soundtrack to elicit 4 emotions: sadness, pleasure, joy, and angry. They tried to classify the self-reported emotions of 5 participant with an artificial neural network.

Unfortunately, the above-mentioned emotion labelling process is unreliable in most cases because it is based on self-reported questionnaires. This issue also has been pointed in some earlier works. For example, in the work of Nagy et al. [4]. They acquired physiological markers (skin conductance) to gauge the arousal state of subject during video watching.

Contrary to earlier emotion recognition studies, in this work we tried to recognize the effect of songs from various music styles instead of emotions. This type of classification does not require any feedback from subjects. Therefore, the whole classification process is more reliable. This approach has been proposed by Suto and Oniga [8] at first. As EEG reader they applied the Neurosky MindWave single-channel device which is one of the simplest and cheapest EEG sensors on the market. In their experiment, 5 subjects listen 10 songs from different styles such as rock, pop, classic, etc. They supposed that the different kinds of songs will generate distinguishable patterns which can be recognized by artificial neural network. The final outcome of their work was an approximately 30-35 recognition accuracy (depending on the subject) with 3-5 second long windows. In this study, we performed a very similar experiment with a 14-channel EMOTIV EPOC+ consumer-grade EEG headset.

2. Methods

2.1. EMOTIV EPOC+

Brainwave signals were acquired using an EMOTIV EPOC+ 14 channel sensor. Its electrodes (saline sensors) are arranged on the AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 positions according to the international 10-20 system. The device also has 2 reference points at P3 and P4 positions (see Fig. 1).

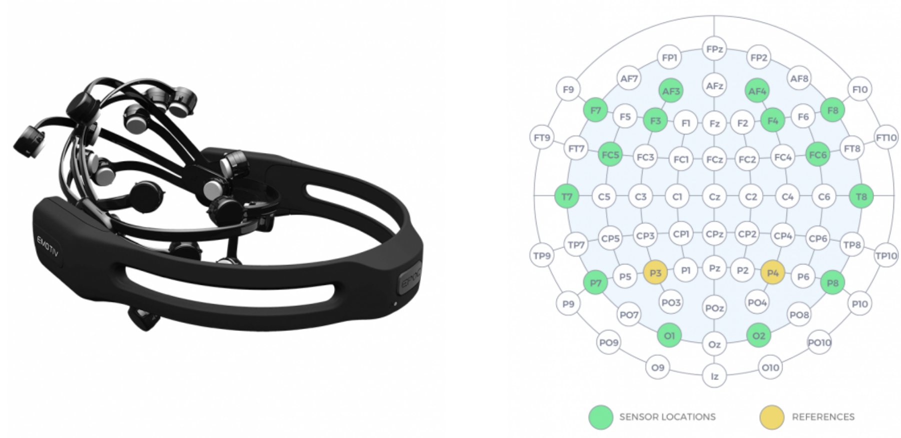

Figure 1: Left: EMOTIV EPOC+ sensor, right: its electrode arrangement on the 10-20 system (source: `www.emotiv.com`)

For EEG recording, the official EMOTIV software has been used. Before data acquisition, electrode impedance should be decreased using saline solution. In the software, users can verify the quality of signals and they can adjust the desired position of all electrodes.

Inside the EEG reader, the sampling method is sequential. The acquired data are digitalized by a 16-bit analog-digital converter (14-bit resolution, 2-bit accuracy) and down-sampled to 128 Hz per channel. It incorporates a built-in notch filter (5th order Sinc) at 50-60 Hz. The communication between EPOC and computer takes place via Bluetooth protocol.

2.2. EEG data recording

In our experiment the data have been collected from 3 volunteers (subjects):

1. subject: 29 years old male
2. subject: 29 years old female
3. subject: 21 years old female

All of them claimed that physiologically and psychologically healthy and right-handed. Each subject has been informed about their task and the goal of the experiment.

The conditions and the data acquisition method were similar as in [8, 9]. During EEG data recording, subjects sat on a chair in front of the computer in a silent room. All subjects listened 1-minute long music pieces from 10 popular songs. They kept their eyes closed and remained as motionless as possible. Between music pieces, there was a 1-2 minutes break. In the experiments, we have followed all principles outlined in the Helsinki Declaration (as revised in 2000).

2.3. Frequency bands and noise

EEG signals can be decomposed into 5 well-known frequency bands:

- delta (δ): 0.5 – 4Hz
- theta (θ): 4 – 8Hz
- alpha (α): 8 – 13Hz
- beta (β): 13 – 30Hz
- gamma (γ): $>$30Hz

Spectral band decomposition is generally applied in all fields of EEG research. All bands can be associated with special emotions and music is also influencing the power of particular bands. Sun et al. [6] observed that light, country, jazz, and rock music have different effect on energy intensity of spectral bands.

Almost all EEG analysis require noise removal because interfering noises can be stronger even 10 times than the real EEG signal. Therefore, all interferences coming from eye movement, power line frequency, electrostatic interference, muscular movement, and heart rhythm significantly affect the quality of the EEG signal. This can prevent the classifier algorithm to recognize patterns in the signal properly. Typically, eye movement is the major noise sources. This is the reason why participants had to listen music with closed eyes.

The easiest way to increase the signal-to-noise ratio is digital filtering. We investigated both the whole band (without frequency subdivision) and the individual frequency bands separately. Band separation has been performed by self-defined finite impulse response (FIR) filters with Hamming window [10]. The frequency response of filters can be seen on Fig. 2.

2.4. Classification with artificial neural network

The EEG signal classification with a shallow artificial neural network (ANN) follows the general machine learning chain: data acquisition, segmentation (or windowing), feature extraction, classifier training and classification.

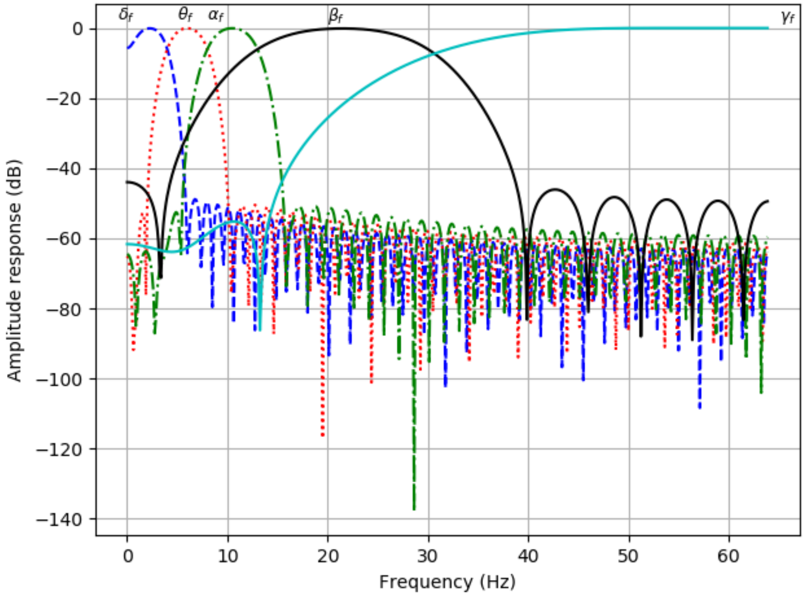

Figure 2: Filters' frequency response

ANN is a strongly parameterized learning algorithm where hyper-parameters such as initial learning rate and weight regularization have significant effect on the model's performance. In this work a two-layer (hidden and output) ANN architecture with 128 hidden neurons has been used according to [7] where the authors compared three shallow ANN architectures on two public datasets (related to human activity recognition). The activation functions on the hidden and output layers were relu (2.1) and softmax (2.2) respectively. In the equations, η_i is the weighted input of the i'th neuron and M is the number of output neurons. The learning algorithm was RMS with no improvement in 20 epochs stop condition and without learning decay. The error function was cross-entropy (2.3) where ω refers to network weights, K is the number of training samples, λ (0.001) is the weight regularization strength, **y** is the target output while **a** is the output activation.

$$\sigma_h(\eta_i) = \max(0, \eta_i) \tag{2.1}$$

$$\sigma_o(\eta_i) = \frac{e^{\eta_i}}{\sum_{j=1}^{M} e^{\eta_j}} \tag{2.2}$$

$$C(\mathbf{y}, \mathbf{a}) = -\sum_{j=1}^{M}[y_j \ln(a_j^L) + (1 - y_j) \ln(1 - a_j^L)] + \frac{\lambda}{2K}\sum_{w}\omega^2. \tag{2.3}$$

In many papers inside machine learning literature, the description about hyper-parameter setup or parameter search is very limited (or missing). However, parameter tuning has a significant effect on the network's performance. The most important parameter of a network is the initial learning rate (α_0). In order to demonstrate the effect of α_0 we applied random parameter search on a 10 base exponential scale (2.4) where exponents come from a uniform distribution ($U(-6, -1)$). The expansion of parameter search (e.g. regularization strength) may cause additional improvement but it was out of the scope of this study.

$$\alpha_0 \in 10^{U(-6,-1)}. \tag{2.4}$$

3. Results

The whole data processing and classification processes have been performed in the Python programming environment. At first, we used the original (full-spectrum) signals from all channels and later their filtered (delta, theta, alpha, beta, and gamma) sub-bands as raw data. An example about the effect of filtering is visible on Fig. 3.

Signals from the 14 channels were aligned line by line into ten data matrices according to the 10 songs. All data matrices went through the windowing process where the window size was 256 samples (2^8). This window size covers 2 seconds wide time interval of the time series. There was no any overlap between windows. From each window the following 12 features have been extracted: *mean, standard deviation, mean absolute deviation, interquartile range, 75th percentile, kurtosis, difference between min and max values, spectral energy, spectral centroid, principal frequency and the first two autoregressive coefficients.* At the end of feature extraction stage all elements have been normalized (mean subtraction and scaling to unit variance). After feature extraction, feature vectors have been randomly divided into 20% test and 80% training data. Finally, 15% of the training set (also randomly selected) were used for validation.

The results of our experiment can be found in Table 1. In the table we can see the highest recognition accuracy (corretly recognized test samples divided by all test semples) of the ANN after random initial learning rate search (100 trials).

4. Conclusion

In this paper, we presented a particular music stimuli recognition experiment with the Emotiv EPOC+ EEG reader which does not require any user feedback. The motivation of this experiment came from [8], where the authors used a Neurosky Mindwave EEG headset for the same task. Compared to their results, we got

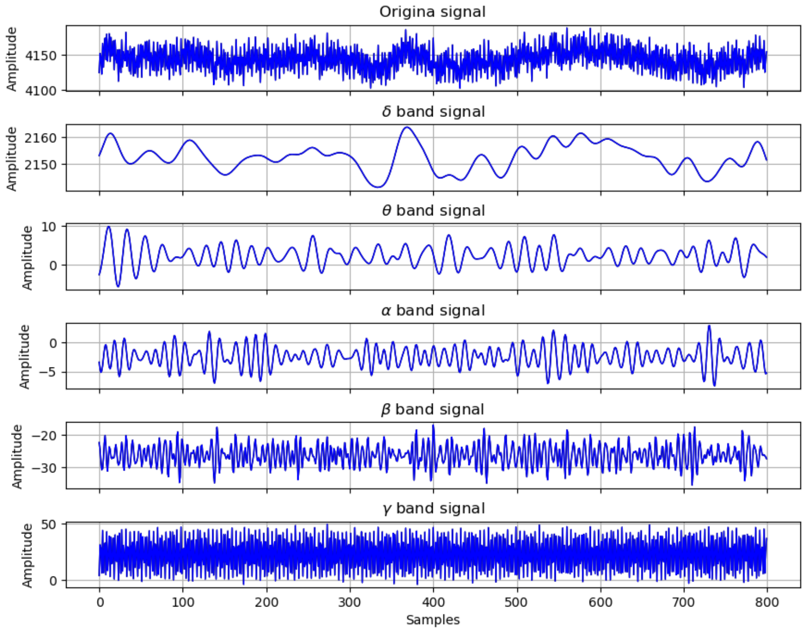

Figure 3: The effect of filtering

Signals	Sub.1	Sub.2	Sub.3
δ band	41.8%	48.3%	35.1%
θ band	34.5%	51.6%	36.8%
α band	47.2%	51.6%	35.1%
β band	47.3%	61.6%	54.4%
γ band	43.6%	**71.6%**	**78.9%**
original signal	**56.3%**	66.6%	68.4%

Table 1: Highest recognition accuracies of the ANN

much better results. Although, on the data of subject 1 the highest recognition accuracy was only 56.3%, in the case of subject 2 and 3 recognition rates were 71.6% and 78.9% respectively. Another interesting difference between our results and [8] lies in the effect of frequency subdivision. In [8] frequency decomposition reduced the recognition accuracy against the unfiltered signal and the classifier was more efficient with wider bands. The latter statement is also met in our work but in the case of subject 2 and 3, frequency decomposition caused significant classification performance improvement. The explanation of this phenomenon requires further investigation. Our results demonstrate that, music has a significant effect on our brain waves but this effect is strongly depending on the observed person. In order to investigate the influence of songs, we have to observe the confusion matrix of the classifier. As an example, the confusion matrix on the data of subject 2 can be seen on Fig. 4.

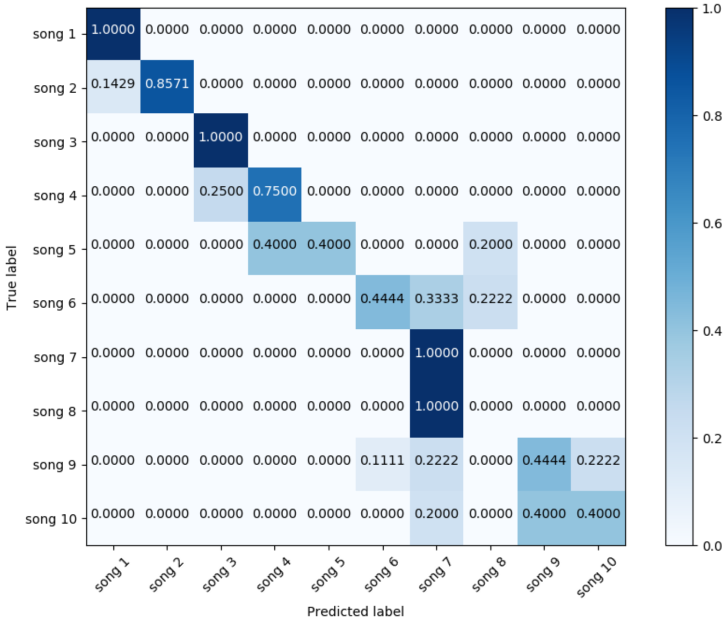

Figure 4: Confusion matrix of subject 2

In the confusion matrix, well visible that the 1., 3., and 7. songs were classified with 100% accuracy while the 8. song (Clint Mansell - Lux aeterna) was totally misclassified. From the individual song classification rates, conclusions can be drawn about the effect of songs.

Our preliminary results, reported in this paper, reflect the effect of songs on the listener's brainwaves without any feedback. Therefore, it would be a majorly enhance in music therapy and music recommendation systems.

Acknowledgements. This work was supported by the construction EFOP-3.6.3-VEKOP-16-2017-00002. The project was supported by the European Union, co-financed by the European Social Fund.

References

[1] Kalaganis, F., Adamos, A.D., Laskaris, N., A consumer BCI for automated music evaluation within a popular on-demand music streaming service, *12th IFIP International Conference on Artificial Intelligence Applications and Innovations*, Thessaloniki, 2015, pp. 429-440.

[2] Lin, Y.P., Wang, C.H., Jeng, S.K., Chen, J.H., Multilayer perceptron for EEG signal classification during listening to emotional music, *TENCON 2007 IEEE Regione 10 Conference*, Taipei, 2007, pp. 1–3.

[3] Meza-Kubo, V., Moran, L.A., Carrillo, I., Galindo, G., Garcia-Canseco, E., Assessing the user experiance of older adults using a neural network trained to recognize emotions from brain signals, *Journal of Biomedical Informatics* Vol. 62 (2016), 202–209.

[4] Nagy, T., Tellez, D., Divák, Á., Lógó, E., Köles, M., Hámornik, B.,Predictiong arousal with machine learning of EEG signals, *5th International Conference on Cognitive Infocommunications*, Vietri Sul Mare, 2014, pp. 137–140.

[5] Ramirez, R., Palencia-Lefler, M., Giraldo, S., Vamvakousis, Z., Musical neurofeedback for treating depression in elderly people, *Frontiers in Neuroscience*, Vol. 9 (2015), 354.

[6] Sun, C. et al., The effect of different types of music on electroencephalogram, *2013 IEEE International Conference on Bioinformatics and Biomedicine*, Shanghai, 2013, pp. 31–37.

[7] Suto, J., Oniga, S., Efficiency investigation of artificial neural network in human activity recognition, *Journal of Ambient Intelligence and Humanized Computing*, Vol. 9 (2018), 1049–1060.

[8] Suto, J., Oniga, S., Music stimuli recognition in electroencephalogram signal, *Elektronika Ir Elektrotechnika*, Vol. 24 (2018), 68–71.

[9] Suto, J., Oniga, S., Pop Sitar, P., Music stimuli recognition from electroencephalogram signal with machine learning, *2018 7th International Conference on Computers Communications and Control (ICCCC)*, Oradea, 2018, pp. 260–264.

[10] Tan, L., Jiang, J., Digital Signal Processing Fundamentals and Applications, 2nd ed., Elsevier, 2013, pp. 161–288.

[11] Wang, X.W., Nie, D., Lu, B.L., Emotional state classification from EEG data using machine learning approach, *Neurocomputing*, Vol. 129 (2014), 94–106.

[12] Yuvaray, R., Murugappan, M., Ibrahim, M.N., Sundaraj, K., Omar, I.M., Mohamad, K., Palaniappan, R., Detection of emotions in Parkinson's disease using higher order spectral features from brain's electrical activity, *Biomedical Signal Processing and Control*, Vol. 14 (2014), 108–116.

[13] Zhang, T., Fruchter, R., Frank, M., Are they paying attention? A model-based method to identify individual's mental states, *Computer*, Vol. 50 (2017), 40–49.