Evaluation of Cepstral Coefficients as Features in EEG-based Recognition of Emotional States

Firgan Feradov[1], Iosif Mporas[2], Todor Ganchev[1]
[1]Technical University of Varna, 1 Studentska str., Varna 9010, Bulgaria
firgan.n.feradov@gmail.com, tganchev@ieee.org

[2]University of Hertfordshire, College Lane Campus, Hatfield AL10 9AB, UK
i.mporas@herts.ac.uk

Abstract. The study of physiological signalsand the evaluation of their featuresare of great importance for theautomated emotion detection, as these are directly connected with the successful modelling and classification of the states of interest. In the presented work, we present an evaluation of the appropriateness of LFCC and the logarithmic energy of signals as features for automated recognition of negative emotional states in terms of recognition accuracy. In particular, three sets of features are compared – featurescomputed afterframe-level segmentationof the signal;featurescomputed after averaging of frame level descriptors; and features extracted froman entire EEG recording. The performance of the extracted features is evaluated using C4.5 classifier for 10, 15, 20, 30, 45, and 60 filters.

Keywords: LFCC features, cepstral coefficients; emotion recognition; EEG

1 Introduction

As the rapid development of computers and smart devices continues and their use is becoming ever more present in the daily life the interest towards the creation of intuitive interfaces and control methods also increases drastically. Affective computing is one of the most prominent directions in this research field. The goal of affective computing is to allow the devices to detect and distinguish different emotional states and adjust the system's behavior in accordance to the user's mood.

There are a number of different approaches to the task of affective computing. One of the major differences in these methods is the modality used for emotion detection – voice recordings, facial features, or physiological signals, as well as some other approaches that exist (body gestures, from written text and others). The detection of emotional states using physiological involves the use of different biological data such as the use of galvanic skin resistance, ECG, EEG and combinations of these and other modalities for detection of emotional states. In recent years, as the recording of EEG signals has become more cost-efficient and accessible the studies focusing on the detection and recognition of affective states from brain activity have increased considerably.

The detection of emotions from EEG signals has a number of advantages. One of its main strengths is that in comparison to the cases where voice or facial expressions are used it is not necessary for the classified affective states to have a strong visible manifestation. This allows for the detection and classification of a wide range and level of emotions. However there are also a number of different problems regarding the use of EEG data, among which are the high dimensionality of the data, the technical difficulties in recording and preprocessing it and the fact that the EEG recording depicts a superposition of the entire brain activity.

The fact that often-different signals, recorded by separate electrodes, contain information, which is relevant to classified emotional state, shows that the use of descriptive features may help solving many of the issues of EEG signal classification. The choice of features does not only affect the accuracy of the classifier, but it is also crucial to the feasibility of the proposed method.One such approach is using Linear Frequency Cepstral Coefficients (LFCC). Studies [1,2] show that some frequency bands in the EEG signal contain information about a person's emotional states. Due to this fact, the cepstral coefficients can be applied in the analysis of emotional states, as they provide a good way to describe activity in different spectral frequency bands. Cepstral coefficients are traditionally used in audio processing, but their application for affective classification from EEG signals has been very limited. Studies that use this approach [3-5] focus on Mel Frequency Cepstral Coefficients (MFCC). Other methods such as LPCC [6] and LFCC [7] have also been examined. The reported results show that the cepstral coefficients can achieve relatively high classification accuracy, with reported results ranging between 70 to 90%.

In this paper, we present a study on the qualities of the two features - the Linear Frequency Cepstral Coefficients (LFCC) and the logarithmic signal energy - extracted from different representations of an EEG signal. The work is presented in five sections. Section 1 presents the studied problem. In Section 2, the feature extraction process is presented. Section 3 discusses the data used for the evaluation and experimental setup. In Section 4, the acquired experimental results are presented and discussed. In Section 5,conclusions are made.

2 Feature extraction

In the presented study two different features are evaluated – the linear frequency cepstral coefficients (LFCC) and the logarithmic energy of an EEG signal. Each of the EEG signals used for the evaluation is 63 seconds long and is recorded using 32 surface electrodes. The recorded raw signals have been preprocessed and all artefacts have been removed. Additionally the signals have been filtered using a 0.5 – 45 Hz band-pass filter.

The evaluated features are extracted from three different representations of the EEG signal. In the first two cases, the signal is segmented to smaller parts using a sliding window with fixed length. In the third case, the features are calculated from the entire signal and no reduction or framing is applied.

For the first two cases the signal is segmented into short frames using a sliding window with 1 second length and 75% overlap between two successive frames. Each channel of EEG signal is processed separately. The total number of frames, generated by the sliding window can be calculated as follows:

$$P = fix\left(\frac{N-K+L}{L}\right) \quad (1)$$

where N is the total number of samples, K is the frame size, L is the step size of the window, and the operator fix denotes that the result is rounded toward the smaller integer number.

For the second experimental setup, the generated segments are averaged:

$$P_m = \frac{\sum_{i=1}^{P} x_{ni}}{P}, \; n = 1,2,...,N \quad (2)$$

wherex_{ni} corresponds to the value of the n-th sample value in the frame P, N is the length of the frame and Pm is the resulting averaged frame.

After the preprocessing and preparation procedures are completed a Discrete Fourier Transform (DFT) is performed on the signals. A filter bank, containing a number triangular window functions is applied to the acquired signal spectrum. The filters in the filter bank are defined as:

$$H_i(k) = \begin{cases} 0 & \text{for } k < f_{b_i-1} \\ \dfrac{(\mathrm{k} - f_{b_i-1})}{(f_{b_i} - f_{b_i-1})} & \text{for} f_{b_i-1} \le k \le f_{b_i} \\ \dfrac{(f_{b_{i+1}} - k)}{(f_{b_{i+1}} - f_{b_i})} & \text{for} f_{b_i} \le k \le f_{b_i+1} \\ 0 & \text{for } k < f_{b_i-1} \end{cases} \quad (3)$$

where the index *i* denotes the number of the number of the filter, f_{bi} specifies the boundaries of the filter, which are expressed as positions, dependent on the used sampling frequency. The index k corresponds to the coefficients of the N-point DFT. The filters are used for acquiring frequency sub-bands of the signal spectrum from which the logarithmic energy (F1) are the LFCC (F2) calculated.

$$S_i = \log_{10}\left(\sum_{k=0}^{N-1} |\mathrm{S}(k)|^2 \cdot H_i(k)\right),\ i = 1,2,\dots,M, \quad (4)$$

where Si is the output of the i-th filter, |S(k)|2 is the power spectrum of the band, N is the DFT size, Hi denotes the i-th filter of the filter bank and M is the total number of filters. Using the obtained Si values the LFCC are calculated by performing decorretion of the filter bank outputs via discrete cosine transform (DCT):

$$LFCC(r) = \sum_{i=1}^{M} S_i\left(\frac{r(i-0.5)}{B}\right),\ r = 0,1,\dots,R-1. \quad (5)$$

where r is the LFCC index, and R ≤ M is the total number of unique LFCC that can be computed.

3 Experimental setup

The performed evaluation experiments were carried out using EEG recordings taken from the DEAP database [8]. DEAP (Database for Emotion Analysis using Physiological signal) is a freely distributed database, containing a number of various types of physiological recordings as EEG, ECG, EMG and others. These physiological recordings are taken from 32 participants who are observing musical videos of songs, varying in style and genre. As a result each of the recordings is rated by the participant based on his emotional response to the shown clips.

The data provided in the dataset is rated using five parameters, namely valance, arousal, dominance, liking and familiarity. Valance, arousal and dominance are parameters that are widely used in tasks regarding the classification and recognition of emotional and affective states and denote the quality of the experienced emotion (positive or negative), the strength of the emotional response and his ability to maintain control over his actions. The two additional parameters, liking and familiarity, included in the dataset, measure the level to which the participant enjoys the song and degree to which he is familiar with the song prior to the test. All of the parameters have range from 1 to 9, where 1 is the lowest rank and 9 is the highest rank. The familiarity rating provides the only exception to this ranking system – in this case the range is from 1 to 5 with 1 being the lowest score and 5 being the highest.

In the presented evaluation the EEG data was split in two groups – negative and others. The separation of the data was performed based on the provided liking ratings. Recordings with liking rating lower than 4 were tagged as “negative”, while recording with liking rating higher than 4 were tagged as positive. The separation was performed on the data of each participant and the results of the class assignment were evaluated. The data of the participants for which the negatively tagged instances comprised less than 20% of the total amount of data were discarded. This pruning process led to the reduction of the number of participants used in the feature evaluation from 32 to 24. In Table 1 we showthe percentage of recordings tagged as "negative" in the DEAP database.

Table 1. The percentage of recordings tagged as "negative" in the DEAP database. Here Pxx stands for the participant number.

P2	**P4**	**P5**	**P6**	**P11**	**P12**	**P13**	**P14**
30%	*40%*	*27.5%*	*20%*	*45%*	*40%*	*22.5%*	*22.5%*
P15	**P16**	**P17**	**P19**	**P20**	**P21**	**P22**	**P23**
27.5%	*55%*	*27.5%*	*27.5%*	*22.5%*	*70%*	*67.5%*	*42.5%*
P24	**P25**	**P26**	**P28**	**P29**	**P30**	**P31**	**P32**
20%	*32.5%*	*22.5%*	*32.5%*	*35%*	*45%*	*22.5%*	*45%*

After the feature extraction process, using the extracted features, the experimental data of each participant is constructed.First the features are used to create feature vectors. The extracted features are grouped based on the channel from which they are calculated from. This way after the processing a feature vector is generated from each channel generates, with a total of 32 feature vectors for each EEG recording. Given that all of the 40 trials, provided in the DEAP database are used for the evaluation, 1280 feature vectors are calculated for each participant. The length of the feature vectors varies depending on the method used to calculate the features. For the cases where signal segmentation is applied the length of the vector is:

$$L_{fv} = P.\mathrm{F} \tag{6}$$

where P is the number of frames and F is the number of filters. For the cases where the features are extracted from an averaged frame or the whole signal the length of the feature vector is equal to the number of filters. Due to this difference the in the length of the feature vectors the data extracted from the framed signal is 245 times larger in size compared to the other two cases.

After the experimental data is preparedthe Weka tool is used to evaluate the qualities of the extracted features for each participant. Weka is a data mining software, created using Java programing language which incorporates a number of different machine learning and data mining algorithms. The program allows the direct use of the provided algorithms trough command lines or incorporation in other Java-based programs.

For the evaluation of the features we use the C4.5 decision tree algorithm provided in Weka. This classification algorithm is chosen due to its fast classification times. Before the experimental evaluation is performed the parameters of the classifier are tuned using a series of test. The performed grid search focuses on the parameters "C", which denotes the confidence factor used for pruning, and "M" which specifies the number of instances per leaf. The parameter combination providing best performance is used during all of the experimental feature evaluations.

4 Results

The presented experimental evaluation is carried out using two sets of features. The first set linear frequency cepstral coefficients are used as features, while in the second the used feature is the logarithmic energy of the signal. The features are extracted using three different signal processing methods. In the first method the signal from each channel was segmented using a sliding window with a fixed overlap. The features were calculated from each generated segment and were used to construct the classification data. In the second method the frames from the different channels are averaged and the features are extracted from the acquired averaged segments. In the third methods no segmentation is performed and the features are extracted directly from the channels signals. For each of these cases the features are calculated six times, with a different number of filters used for each calculation. The experiments are carried out with features extracted using 10, 15, 20, 30, 45 and 60 filters.

The results of the performed evaluation are presented in tables 2 and 3. In table 2 the results of the LFCC tests are presented, while in 3 the results achieved using logarithmic energy of the signal as a feature are given. In the first column of the tables the number of filters used for the experiment are displayed. The second column contains the results achieved using features, extracted from the segmented signal (FS – Frames of Signal). In the third column the classification results of the features extracted from the averaged signal frames (MFS -Mean Frame of Signal) is given. In the last column contains the results achieved using features extracted from the signal, when no segmentation is performed (WS – Whole Signal).

Table 2. Experimental resultsacquired using LFCC as feature

# of filters	FS	MFS	WS
F10	86.3 %	75.3 %	74.7 %
F15	86.2 %	76.0 %	75.6 %
F20	85.3 %	76.3 %	76.3 %
F30	85.1 %	76.5 %	77.9 %
F45	84.4 %	76.3 %	78.7 %
F60	83.2 %	75.1 %	79.6 %

Table 3. Experimental results acquired using logarithmic energy as feature

# of filters	FS	MFS	WS
F10	86.4 %	70.7 %	74.4 %
F15	86.9 %	72.2 %	77.1 %
F20	86.3 %	72.6 %	79.2 %
F30	86.6 %	73.8 %	83.1 %
F45	85.7 %	74.6 %	85.3 %
F60	86.2 %	74.8 %	86.6 %

The lowest classification accuracy achieved after during the experimental evaluation was 70.7% and was achieved with logarithmic energy extracted using 10 filtersfrom the averaged signal frames. The highest accuracy was achieved again with logarithmic energy by using 15 filters extracted on the framed signals.

By examining the results it can be seen that the performance of the LFCC and logarithmic energy is very close in the cases where the features are extracted from the segmented EEG signal. In those cases we can also observe the highest accuracies. This is due to the fact that the scanning of the EEG signal using a sliding window allows the extraction of information about the temporal changes in the signal. In the second case this information is lost after the averaging of the frames which leads to a drop in accuracy. In most of these cases it can be seen that the increase of the number of filters does not affect the classification accuracy, with the fluctuations in the results being in the margin of error. Exception to these observations can be seen in the experiments using logarithmic energy extracted from the averaged signal, where the increase of classification accuracy with the increase of the number of filters. The same patterns can be seen in the experiments based on features extracted from the whole signal, were the increase of the number of filters used has a significant effect on the classification results. In the case where logarithmic energy is used the increases of the filters causes12.2% increase in accuracy, reaching 86.6%, which is comparable to the results achieved using features extracted from the framed signals.

5 Conclusion

In the presented paper, we have performed an evaluation of the LFCC and the logarithmic energy of an EEG signal as features for the purposes of emotion classification. The features were extracted using three different approaches – from signal frames acquired using a sliding window, from an averaged frame for each channel and by direct calculation from the signal channel, without any segmentation being performed.

The conducted experimental evaluation has shown that the best classification accuracy for both the cepstral coefficients and logarithmic energy was achieved using the features extracted from the framed signal, followed by the features extracted from the whole signal. The lowest classification accuracy was achieved using the features extracted from the mean signal frames.

The most crucial difference between the three methods is the amount of generated feature data. The segmentation of the signal into frames resulted in the generation of large numbers of features, which required the creation of feature super-vectors for the classification process. This led to substantial difference in the model creation and classification time between the different approaches.

While the second two approaches generated the same amount of features and the classification time required was identical, the features extracted from the whole signal achieved higher accuracy. This difference was considerable in the cases, were logarithmic energy was used. When the computational time and power needed to calculated the frames is taken into account we can conclude that the using the features extracted from the whole EEG signal provides best performance when accuracy and cost efficiency are considered.

Future research will be focused on methods for further reduction of the data size and reducing classification times while maintaining and increasing the achieved classification accuracy. The final goal of the conducted research is the creation of robust features and models, applicable to real-time emotion detection.

References

1. Ahern, G.L. and Schwartz, G.E., 1985. Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis.Neuropsychologia, 23(6), pp.745-755.http://doi.org/10.1016/0028-3932(85)90081-8
2. Rowland, N., Meile, M.J. and Nicolaidis, S., 1985. EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. Science, 228(4700), pp.750-752.
3. M. Othman, A. Wahab and R. Khosrowabadi, "MFCC for robust emotion detection using EEG," 2009 IEEE 9th Malaysia International Conference on Communications (MICC), Kuala Lumpur, 2009, pp. 98-101.doi: 10.1109/MICC.2009.5431473
4. Wahab, A., Kamaruddin, N., Palaniappan, L.K., Li, M. and Khosrowabadi, R., 2010.“EEG signals for emotion recognition. Journal of Computational Methods in Sciences and Engineering”, 10(s1), pp.1-11. DOI: 10.3233/JCM-2010-0263
5. D. Handayani, H. Yaacob, A. Wahab and I. F. T. Alshaikli, "Statistical Approach for a Complex Emotion Recognition Based on EEG Features," 2015 4th International Conference on Advanced Computer Science Applications and Technologies (ACSAT), Kuala Lumpur, 2015, pp. 202-207. doi: 10.1109/ACSAT.2015.54
6. Pazhanirajan, S., and P. Dhanalakshmi. "EEG signal classification using linear predictive cepstral coefficient features." International Journal of Computer Applications 73.1 (2013).
7. Feradov F. “Study of the quality of Linear Frequency CepstralCoefficients for automated recognition of negative emotional states from EEG signals”, Scientific works of the Union of Scientist in Bulgaria – Plovdiv, Series G. Medicine, Pharmacy and Dental Medicine, Vol XIX, ISSN 1311-9427, June 2016, pp 106-109
8. S. Koelstra*et al*., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," in *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, Jan.-March 2012. doi: 10.1109/T-AFFC.2011.15