

Emotion Assessment by Variability-Based Ranking of Coherence Features from EEG

Iván De La Pava^(✉), Andres Álvarez-Meza, and Alvaro-Angel Orozco

Automatic Research Group, Faculty of Engineering,
Universidad Tecnológica de Pereira, Pereira, Colombia
ide@utp.edu.co

Abstract. The automatic assessment of emotional states has important applications in human-computer interfaces and marketing. Several approaches use a dimensional characterization of emotional states along with features extracted from physiological signals to classify emotions elicited from complex audiovisual stimuli; however, the classification accuracy remains low. Here, we develop an emotion assessment approach using a variability-based ranking scheme to reveal relevant coherence features from electroencephalography (EEG) signals. Our method achieves higher classification accuracies than comparable state-of-the-art methods and almost matches the performance of multimodal strategies that require information from several physiological signals.

Keywords: Emotion assessment · Electroencephalography
Coherence features · Relevance analysis

1 Introduction

The automatic recognition of emotional states is an active area of research due to its wide range of possible applications, including human-computer interfaces (HCI), targeted publicity, automatic video tagging, among others [2, 8]. Usually, emotional states are studied using either a discrete or a dimensional representation. The discrete representation associates them to a series of specific emotions, e.g. joy, anger, fear, sadness, etc. On the other hand, the dimensional approach uses a set of latent dimensions that describe each emotional state regarding the physiological responses that emotions induce. The dimensional representation has the advantage of allowing a broad range of emotions to be described as a combination of the basic dimensions. Commonly, the arousal/valence dimensional spaces are considered, where each emotion is represented in terms of the active or passive response to a stimulus (arousal dimension), and the positive or negative response to the same stimulus (valence dimension) [5].

Once the emotional state representation has been established, the type of data from which those states can be assessed must be selected. Both audiovisual (facial expressions and speech patterns) and physiological data have been used for this purpose [2]. The latter have the advantage of not being regulated by

the subject. In particular, the information provided by electroencephalography (EEG) signals has well-established connections to cognitive processes and emotional states [3, 4]. Nonetheless, the use of EEG signals for emotion recognition is far from being straightforward, since complex spatiotemporal relationships between EEG channels have to be characterized to capture patient and stimuli specific variations associated with each emotional state. In [10] power spectral density features extracted from EEG are used for emotional state classification in the arousal/valence dimensional space. Others have proposed more complex feature extraction schemes based on wavelet decomposition [2], brain connectivity measures and graph theory [4, 6], or a characterization by spatially compact regions of interest [8]. In [11] a multimodal approach using EEG and other six physiological signals, along with an ensemble deep learning classifier, was used to recognize binary arousal or valence states. However, the accuracies achieved by these approaches when complex stimuli, such as music videos, are used to elicit the emotional response remain relatively low, and there is still uncertainty about which features of the EEG signal are the most relevant.

Here, we introduce a patient dependent approach for emotional state classification in the arousal/valence dimensional space, using a functional brain connectivity based characterization of EEG signals. The functional brain connectivity measure, termed coherence, is used to extract frequency and spatiotemporal dependent features from EEG. In turn, these features are ranked according to a variability-based relevance analysis, and input to a K-nearest neighbor classifier. Obtained results show that our method outperforms the state-of-the-art techniques concerning the classification accuracy in a public database while using a simple classifier and features extracted only from EEG signals. Also, it almost matches the performance of multimodal strategies that require information from several physiological signals. The remainder of the paper is organized as follows: Sect. 2 introduces the theoretical foundations, Sect. 3 describes the experiments and obtained results, and Sect. 4 presents the conclusions.

2 Variability-Based Ranking of the Coherence Features

Let $\Psi = \{\mathbf{X}_n \in \mathbb{R}^{C \times M}\}_{n=1}^N$ be an EEG set holding N trials of an emotion elicitation experiment with C channels and M samples. Besides, let $\mathbf{Y} \in \{-1, 1\}^N$ be a label set for a particular experiment; where the n -th element y_n corresponds to the emotion dimension class obtained for trial \mathbf{X}_n . Our goal is to infer the class label y_n from a set of features extracted from \mathbf{X}_n . Thus, we use a functional brain connectivity analysis, termed magnitude square coherence or coherence, to code spatiotemporal and frequency dependencies among EEG channels as follows: let $\mathbf{x}_c, \mathbf{x}_{c'} \in \mathbb{R}^M$ be a pair of simultaneously recorded EEG signals belonging to \mathbf{X}_n , their coherence $\gamma_{cc'}(f) \in [0, 1]$ at frequency f is given by:

$$\gamma_{cc'}(f) = \frac{|S_{cc'}(f)|^2}{S_{cc}(f)S_{c'c'}(f)}, \quad (1)$$

where $S_{cc'}(f) \in \mathbb{C}$ is the cross-spectrum between \mathbf{x}_c and $\mathbf{x}_{c'}$, and $S_{cc}(f), S_{c'c'}(f) \in \mathbb{C}$ are the auto-spectral densities of \mathbf{x}_c and $\mathbf{x}_{c'}$, respectively. Namely,

the cross-spectrum is the Fourier transform of the cross-correlation function between the two signals. Analogous relationships hold for the auto-spectral densities. $\gamma_{cc'}(f)$ is a linear measure of the relationship stability between \mathbf{x}_c and $\mathbf{x}_{c'}$ with respect to power asymmetry and phase behavior [9]. Thus, a matrix $\mathbf{\Gamma}_n \in \mathbb{R}^{P \times C(\frac{C-1}{2})}$ holding the coherence values for a range of frequencies $\mathbf{f} \in \mathbb{R}^P$ between all pairwise, non-repeating EEG channel combinations characterizes the trial \mathbf{X}_n regarding linear functional relationships in the frequency and spatial domains. The latter because of the spatial location of the EEG electrode associated with each channel. Temporal information can also be included in the coherence based characterization by splinting each channel \mathbf{x}_c into Q segments through a windowing procedure, using a window of length $L < M$, so that for each channel there is a set $\{\mathbf{z}_c^j \in \mathbb{R}^L\}_{j=1}^Q$. The coherence analysis is then performed over each element of a set of matrices $\{\mathbf{Z}_n^j \in \mathbb{R}^{C \times L}\}_{j=1}^Q$, holding row vectors \mathbf{z}_c^j . The result is a set of coherence matrices $\mathbf{\zeta}_n = \{\mathbf{\Gamma}_n^j \in \mathbb{R}^{P \times C(\frac{C-1}{2})}\}_{j=1}^Q$ which could be used directly to infer y_n . However, the high dimensionality of $\mathbf{\zeta}_n$ poses a discrimination hurdle, and the need to rank the features contained in $\mathbf{\zeta}_n$ arises, so that only the most relevant ones are used in further classification stages.

In this sense, a variability dependent relevance analysis, based on Principal Component Analysis (PCA), is performed as follows: vector concatenation is applied to $\mathbf{\zeta}_n$ to yield a vector $\boldsymbol{\lambda}_n \in \mathbb{R}^D$, with $D = P \times C(\frac{C-1}{2}) \times Q$. Then, the N coherence vectors $\boldsymbol{\lambda}_n$, corresponding to each trial, are stacked to form a matrix $\mathbf{\Lambda} \in \mathbb{R}^{N \times D}$ with the coherence information of the entire EEG dataset Ψ . Afterwards, a relevance index vector $\boldsymbol{\rho} \in \mathbb{R}^D$ is computed as follows:

$$\boldsymbol{\rho} = \mathbb{E}(|\rho_d \boldsymbol{\alpha}_d| : \forall d \in D' \leq D), \quad (2)$$

where $\rho_d \in \mathbb{R}^+$ and $\boldsymbol{\alpha}_d \in \mathbb{R}^D$ are the eigenvalues and eigenvectors of the covariance matrix $\mathbf{\Lambda}^\top \mathbf{\Lambda} / D$, and D' depends on the percentage of variance retained from the input data [1]. Under the assumption that emotions are better assessed in the frequency domain [4, 10], the relevance vector $\boldsymbol{\rho}$ is averaged over the frequency using knowledge about the structure of $\boldsymbol{\lambda}_n$. The average vector $\bar{\boldsymbol{\rho}} \in \mathbb{R}^D$ can then be used to rank $\boldsymbol{\lambda}_n$ according to the most discriminant features, that is, the frequencies that present the highest variability in the coherence values. Therefore, only the $p < D$ features $\boldsymbol{\lambda}_n^p \in \mathbb{R}^H$, with $H = p \times C(\frac{C-1}{2}) \times Q$, corresponding to the most relevant frequencies as indicated by the p largest values of $\bar{\boldsymbol{\rho}}$, are selected to discriminate emotional states.

3 Experiments and Results

To test the proposed method for emotional state classification, we used the pre-processed version of the Database for Emotion Analysis Using Physiological Signals (DEAP) [5]. This database consists of EEG records taken from 32 subjects while performing 40 trials of an emotion elicitation experiment. In each experiment the subjects were exposed to a music video, then they rated their response to the each video in two scales, from 1 to 9, in the valence and arousal

emotional dimensions. The valence scale ranged from unpleasant to pleasant, while the arousal scale ranged from inactive or calm to active or excited. The EEG data were acquired using a 32 channel BioSemi ActiveTwo system. The preprocessed dataset underwent artifact removal, a frequency down-sampling to 128 Hz, and a bandpass filtering from 4–45 Hz. Besides, the data was averaged to the common reference and segmented into 60 s trials. Since our goal is to use EEG data of a particular subject to infer his or her emotional response to a stimulus, we recast the emotion state assessment task as two binary classification problems. The rating scales for both valence and arousal were divided into low (from 1 to 5) and high (from 5 to 9) levels, and given class labels -1 and 1 , respectively. To characterize the EEG data corresponding to each trial, the coherence was computed as explained in section Sect. 2, using 12 s long Hamming window with a 50% overlap to segment the signals and add temporal variation to the coherence data [2]. Only the coherence values for frequencies between 5 Hz and 45 Hz were considered, because of abnormally high variance present beyond the cutoff frequencies of the bandpass filter applied to the data. Then, to rank the coherence features in the frequency domain, the variability-based analysis described in section Sect. 2 was performed, setting the percentage of retained variance to 98%. Next, the features ranked according to the average relevance vector $\bar{\rho} \in \mathbb{R}^P$, $P = 161$, were used as inputs, in a progressive and accumulative fashion, to a classification algorithm. The classification procedure was carried out using a Gaussian similarity-based K-nearest neighbor classifier. The number of nearest neighbors of the classifier was selected for each subject from the range $K = \{1, 3, 5, 7, 9, 11\}$ as the one producing the highest classification accuracy. We employed a nested cross-validation scheme of 10 repetitions, where 70% of trials were used as the training set and the remaining 30% as the test set.

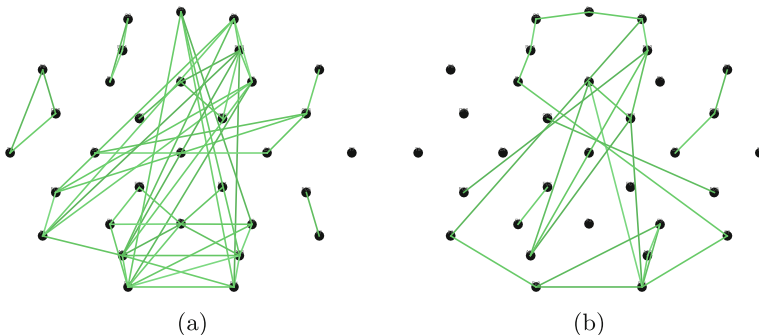


Fig. 1. Graphical representation of the coherence between EEG channels for a DEAP subject. The plots show the coherence for the trials rated as eliciting (a) minimum valence, and (b) maximum valence in the original rating scale ranging from 1 to 9. (Coherence plots obtained with the Matlab toolbox HERMES [7]).

Figure 1 shows a graphical representation of the coherence between EEG channels for the trials rated as eliciting minimum valence (Fig. 1(a)), and maximum valence (Fig. 1(b)) in one of the subjects of the database. The nodes of the

plots represent the spatial distribution of the 32 EEG channels associated with the positions of the corresponding electrodes on the scalp (the uppermost nodes correspond to the electrodes located closer to the anterior part of the head). The lines joining the nodes represent coherence values, at a given frequency, between the joined nodes or channel pairs larger than a threshold level. Our method aims to exploit the difference in brain functional connectivity generated by different emotional states, as measured by the coherence, to identify those states from EEG data.

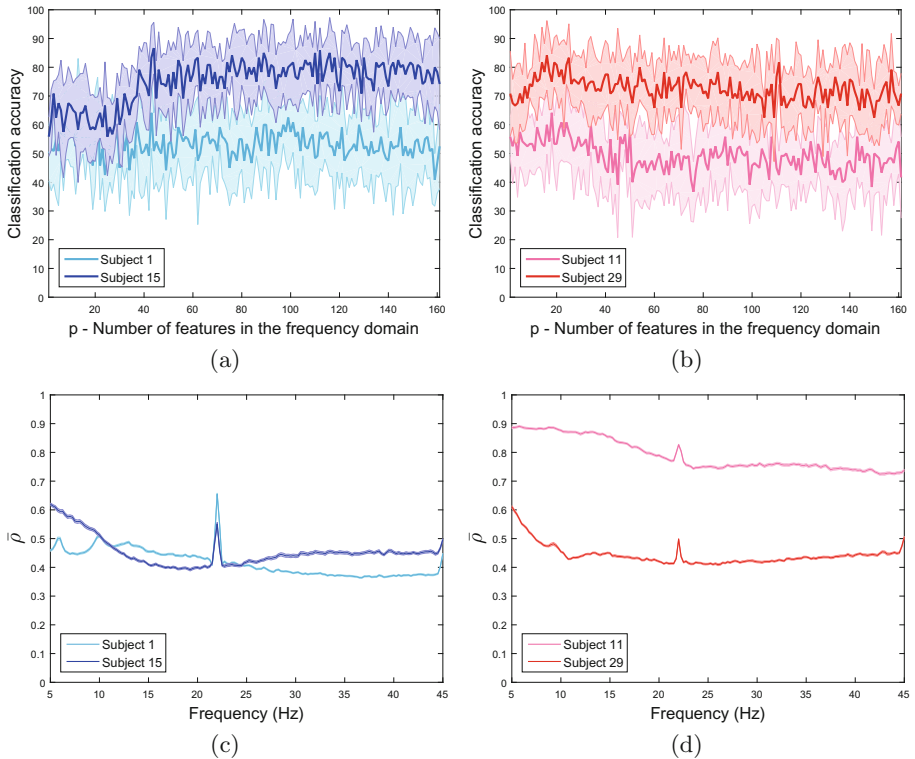


Fig. 2. Top row: average classification accuracy for two different subjects as a function of the number of relevant coherence features p for (a) valence and (b) arousal. Bottom row: average relevance vectors $\bar{\rho}$ used to rank the coherence features, the vectors in (c) were used to rank the features in (a), and analogously for (d) and (b).

To take advantage of the wealth of information provided by the coherence, while avoiding the problems posed by the high dimensionality of these data, we introduced a variability based relevance analysis of the coherence features. This analysis allows ranking the features so that they can be input to the classifier in a progressive and accumulative way. Figure 2(a) and (b) show the effects on the classification accuracy, for the valence and arousal emotional dimensions, of using

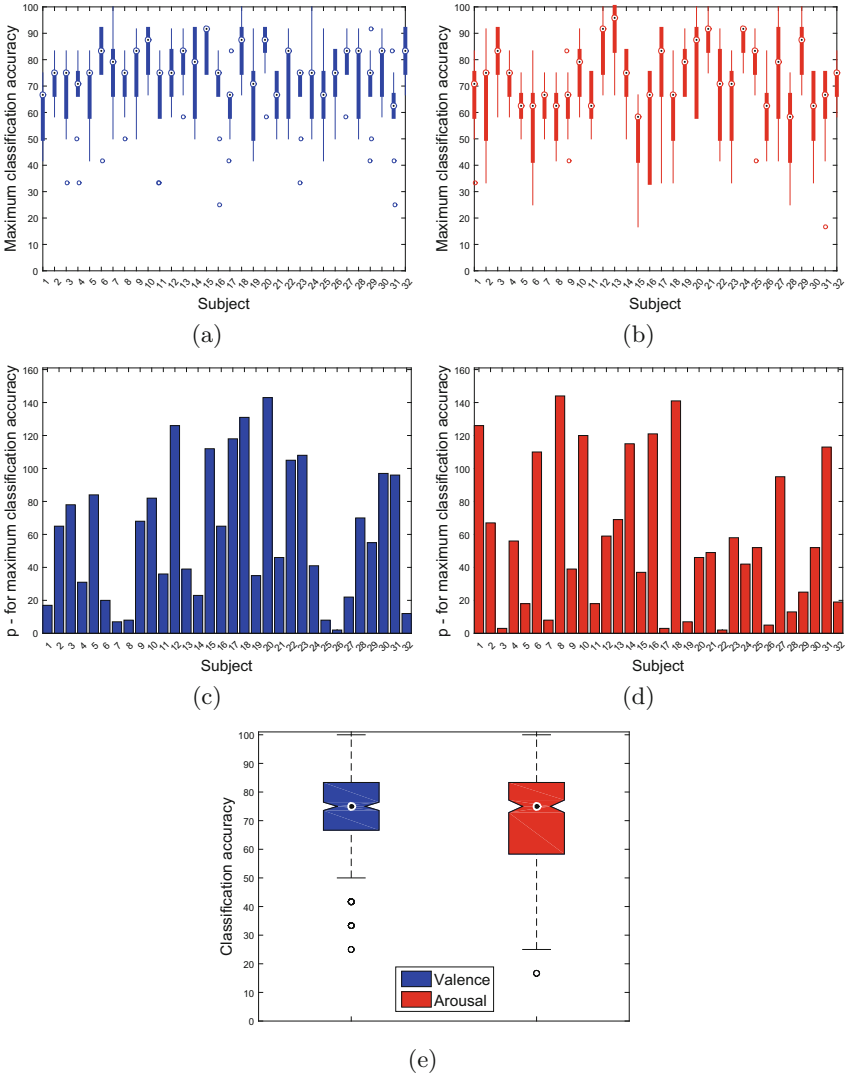


Fig. 3. Maximum classification accuracy distributions for (a) valence, and (b) arousal, holding the number of coherence features in (c) and (d), respectively. (e) shows the maximum valence and arousal classification accuracy distributions for all subjects.

an increasing number p of features as inputs to the classifier. The first feature used corresponds to the most relevant one in the frequency domain according to the relevance vector $\bar{\rho}$. Figure 2(c) and (d) show the average values of $\bar{\rho}$ used to rank the features employed to obtain Fig. 2(a) and (b), respectively. The behavior of $\bar{\rho}$ for each subject is remarkably stable, with small variations for the different folds of the cross-validation scheme. This behavior points to the feasibility of

using a variability index obtained from a training set to rank the coherence features of a new input from the same subject. Larger variations in $\bar{\rho}$ are evident in an inter-subjects comparison. Such comparison also shows that, in average, the most relevant features correspond to the theta (θ : 4–8 Hz) and alpha (α : 8–13 Hz) frequency bands. These frequency bands, and especially the alpha band, have been widely used as indicators of emotional states [4]. The peak midway between the cutoff frequencies of the bandpass filter applied to the data is more likely be the result of variation introduced in the coherence by the filter itself, rather than being a reflection of relevant neurophysiological activity. Another significant result that can be observed in Fig. 2 is that the method achieves its best performance, with respect to the classification accuracy, for a p value that can be lower than the total number of coherence features P (Fig. 2(a) and (b)). Besides, the value of p that produces the maximum classification accuracy varies among subjects and between the emotion dimensions. These results are presented more clearly in Fig. 3.

Figure 3(a) and (b) show the maximum classification accuracies for each subject for the level of valence and arousal, respectively. Figure 3(c) and (d) show the number of coherence features p at which the system reached the maximum accuracies for each subject. Figure 3(c) and (d) highlight the fact that the number of input features required to attain the best performance of the system varies both among subjects and within subjects when either valence or arousal are assessed. When the adequate p is selected, the system achieves median classification accuracies for every subject in the database between 60% and 91% for the valence level, and between 55% and 96% for the arousal level. The overall performance of the system under the same condition stated above is shown in Fig. 3(e). The maximum classification accuracy distributions for all subjects show a median accuracy of 75% for both the valence and arousal levels. Those results put the performance of the proposed method at the top of similar approaches in the state-of-the-art that tackle the emotion assessment problem and test their methods in the DEAP database, and only slightly below the multimodal approach proposed in [11] that uses not only EEG but also other six physiological signals. A comparison with such methods is presented in Table 1.

Table 1. Classification accuracy [%] for all subjects in DEAP.

Approach	Arousal	Valence
Koelstra et al. [5]	62.00	57.60
Soleymani et al. [10]	62.10	50.50
Gupta et al. [4]	65.00	60.00
Padilla-Buritica et al. [8]	52.80	58.60
Daimi et al. [2]	66.90	65.30
Yin et al. (multimodal strategy) [11]	77.19	76.17
Our method	75.00	75.00

4 Conclusions

In this study, we developed a patient-specific method to assess from EEG signals the emotional state, in the arousal/valence dimensional space, elicited by audio-visual stimuli. Our method first characterizes the EEG signals using coherence, a functional brain connectivity measure. Next, the frequency and spatiotemporal temporal information contained in the coherence features is ranked according to a variability analysis based on PCA, under the assumption that those features with the highest degree of variability will be the most discriminant, then the classification is carried out using a K-nearest neighbor classifier. Our approach outperforms most of the state-of-the-art methods that use EEG for emotional state assessment regarding the classification accuracy and obtains results close to those of a multimodal approach that uses data from several physiological signals. As future work, our method can likely be further improved by the utilization of a more robust classifier, e.g., support vector machine, and by the inclusion of directionality in the brain connectivity analysis, e.g., partial directed coherence.

Acknowledgments. This work was supported by the project “Desarrollo de un sistema de apoyo al diagnóstico no invasivo de pacientes con epilepsia fármaco-resistente asociada a displasias corticales cerebrales: método costo-efectivo basado en procesamiento de imágenes de resonancia magnética” with code 1110-744-55778, and author I. De La Pava was supported by the program “Doctorado Nacional en Empresa - Convectoria 758 de 2016”, both funded by Colciencias. The authors would also like to thank Cristian Alejandro Torres Valencia for his valuable contribution to this work.

References

1. Álvarez-Meza, A.M., et al.: Time-series discrimination using feature relevance analysis in motor imagery classification. *Neurocomputing* **151**, 122–129 (2015)
2. Daimi, S.N., et al.: Classification of emotions induced by music videos and correlation with participants rating. *Expert Syst. Appl.* **41**(13), 6057–6065 (2014)
3. Gonuguntla, V., et al.: Identification of emotion associated brain functional network with phase locking value. In: 2016 IEEE 38th Annual International Conference of the EMBC, pp. 4515–4518. IEEE (2016)
4. Gupta, R., et al.: Relevance vector classifier decision fusion and EEG graph-theoretic features for automatic affective state characterization. *Neurocomputing* **174**, 875–884 (2016)
5. Koelstra, S., et al.: Deap: a database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* **3**(1), 18–31 (2012)
6. Lee, Y.Y., Hsieh, S.: Classifying different emotional states by means of EEG-based functional connectivity patterns. *PLoS One* **9**(4), e95415 (2014)
7. Niso, G., et al.: HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity. *Neuroinformatics* **11**(4), 405–434 (2013)
8. Padilla-Buritica, J.I., Martínez-Vargas, J.D., Castellanos-Dominguez, G.: Emotion discrimination using spatially compact regions of interest extracted from imaging EEG activity. *Front. Comput. Neurosci.* **10**, 1–11 (2016)

9. Sakkalis, V.: Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG. *Comput. Biol. Med.* **41**(12), 1110–1117 (2011)
10. Soleymani, M., et al.: Multimodal emotion recognition in response to videos. *IEEE Trans. Affect. Comput.* **3**(2), 211–223 (2012)
11. Yin, Z., et al.: Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. *Comput. Methods Programs Biomed.* **140**, 93–110 (2016)