We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Affective Valence Detection from EEG Signals Using Wrapper Methods

Antonio R. Hidalgo-Muñoz, Míriam M. López,
Isabel M. Santos, Manuel Vázquez-Marrufo,
Elmar W. Lang and Ana M. Tomé

Additional information is available at the end of the chapter

Abstract

In this work, a novel valence recognition system applied to EEG signals is presented. It consists of a feature extraction block followed by a wrapper classification algorithm. The proposed feature extraction method is based on measures of relative energies computed in short-time intervals and certain frequency bands of EEG signal segments time-locked to the stimuli presentation. These measures represent event-related desynchronization/synchronization of underlying brain neural networks. The subsequent feature selection and classification steps comprise a wrapper technique based on two different classification approaches: an ensemble classifier, i.e., a random forest of classification trees and a support vector machine algorithm. Applying a proper importance measure from the classifiers, the feature elimination has been used to identify the most relevant features of the decision making both for intrasubject and intersubject settings, using single trial signals and ensemble averaged signals, respectively. The proposed methodologies allowed us to identify a frontal region and a beta band as the most relevant characteristics, extracted from the electrical brain activity, in order to determine the affective valence elicited by visual stimuli.

Keywords: EEG, random forest, SVM, wrapper method

1. Introduction

During the last decade, information about the emotional state of users has become more and more important in computer-based technologies. Several emotion recognition methods and their applications have been addressed, including facial expression and microexpression recognition, vocal feature recognition and electrophysiology-based systems [1]. More recently,

INTECH
open science | open minds

the integration of emotion forecasting systems in ambient-assistant living paradigms has been considered [2]. Concerning the origin of the signal sources, the used signals can be divided into two categories: those originating from the peripheral nervous system (e.g., heart rate, electromyogram, galvanic skin resistance, etc.) and those originating from the central nervous system (e.g., electroencephalogram (EEG)). Traditionally, EEG-based technology has been used in medical applications but nowadays it is spreading to other areas such as entertainment [3] and brain-computer interfaces (BCI) [4]. With the emergence of wearable and portable devices, a vast amount of digital data are produced and there is an increasing interest in the development of machine-learning software applications using EEG signals. For the efficient manipulation of this high-dimensional data, various soft computing paradigms have been introduced either for feature extraction or pattern recognition tasks. Nevertheless, up to now, as far as authors are aware, few research works have focused on the criteria to select the most relevant features linked to emotions, relying most of the studies on basic statistics.

It is not easy to compare different emotion recognition systems, since they differ in the way emotions are elicited and in the underlying model of emotions (e.g., discrete or dimensional model of emotions) [5]. According to the dimensional model of emotions, psychologists represent emotions in a 2D valence/arousal space [6]. While valence refers to the pleasure or displeasure that a stimulus causes, arousal refers to the alertness level which is elicited by the stimulus (see **Figure 1**). Sometimes an additional category assigned as *neutral* is included, which is represented in the region close to the origin of the 2D valence/arousal space. Some studies concentrate on one of the dimensions of the space such as identifying the arousal intensity or the valence (low/negative versus high/positive) and eventually a third class neutral state. Recently, it was pointed out that data analysis competitions, similar to the brain-computer interfaces community, could encourage the researchers to disseminate and compare their methodologies [7].

Normally, emotions can be elicited by different procedures, for instance by presenting an external stimulus (picture, sound, word, or video), by facing a concrete interaction or situation [8] or by simply asking subjects to imagine different kinds of emotions. Concerning external visual stimuli, one may resort to standard databases such as the international affective picture system (IAPS) collection which is widely used [7, 9] or the DEAP database [10] that also includes some physiological signals recorded during multimedia stimuli presentation. Similar to any other classification system, in physiology-based recognition systems, it is needed to establish which signals will be used to extract relevant features from these input signals and finally to use them for training a classifier. However, as often it occurs in many biomedical data applications, the initial feature vector dimension can be very large in comparison to the number of examples to train (and evaluate) the classifier.

In this work, we prove the suitability of incorporating a wrapper strategy for feature elimination to improve the classification accuracy and to identify the most relevant EEG features (according to the standard 10/20 system). We propose it by using the spectral features related to EEG synchronization, which has never been applied before for similar purposes. Two learning algorithms integrating the classification block are compared: random forest and support vector machine (SVM). In addition, our automatic valence recognition system has been tested

both in intra and intersubject modalities, whose input signals are single trials (segments of signal after the stimulus presentation) of only one participant and ensemble averaged signals computed for each stimulus category and every participant, respectively.
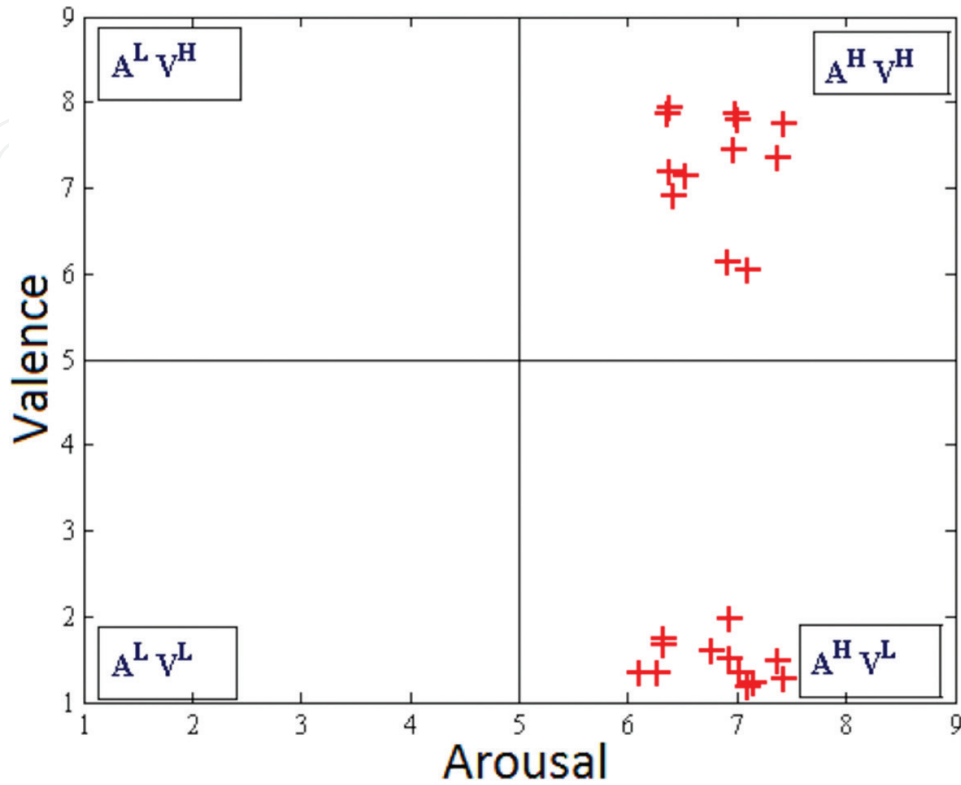

Figure 1. Ratings of the pictures selected from international affective picture system for carrying out the experiment. L: low rating; H: high rating.

2. Related work

The following subsections review some examples of machine learning approaches to affective computing and brain cognitive works where time-domain and frequency-domain signal features are related to the processing of emotions.

2.1. Classification systems and emotion

The pioneering work of Picard [11] on affective computing reports a recognition rate of 81%, achieved by collecting blood pressure, skin conductance and respiration information from one person during several weeks. The subject, an experienced actor, tried to express eight affective states with the aid of a computer-controlled prompting system. In Ref. [12], using the IAPS data set as stimulus repertoire, peripheral biological signals were collected from a single person during several days and at different times of the day. By using a neural network classifier, they considered that the estimation of the valence value (63.8%) is a much harder task than the estimation of arousal (89.3%). In Ref. [13], a study with 50 participants, aged from 7 to

8 years old, is presented. The visual stimulation with the IAPS data set was considered insufficient, hence they proposed a sophisticated scenario to elicit emotions and only peripheral biological signals were recorded and the measured features were the input of a classification scheme based on an SVM. The results showed accuracies of 78.4% and 61% for three and four different categories of emotions, respectively.

In Ref. [14], also by means of the IAPS repository, three emotional states were induced in five male participants: *pleasant*, *neutral* and *unpleasant*. They obtained, using SVMs, an accuracy of 66.7% for these three classes of emotion, solely based on features extracted from EEG signals. A similar strategy was followed by Macas [15], where the EEG data were collected from 23 subjects during an affective picture stimulus presentation to induce four emotional states in arousal/valence space. The automatic recognition of the individual emotional states was performed with a Bayes classifier. The mean accuracy of the individual classification was about 75%.

In Ref. [16], four emotional categories of the arousal/valence space were considered and the EEG was recorded from 28 participants. The ensemble average signals were computed for each stimulus category and person. Several characteristics (peaks and latencies) as well as frequency-related features (event-related synchronization) were measured on a signal ensemble encompassing three channels located along the anterior-posterior line. Then, a classifier (a decision tree, *C*4.5 algorithm) was applied to the set of features to identify the affective state. An average accuracy of 77.7% was reported.

In Ref. [17], through a series of projections of facial expression images, emotions were elicited. EEG signals were collected from 16 healthy subjects using only three frontal EEG channels. In Ref. [18], four different classifiers (quadratic discriminant analysis (QDA), k-nearest neighbor (KNN), Mahalanobis distance and SVMs) were implemented in order to accomplish the emotion recognition. For the single channel case, the best results were obtained by the QDA (62.3% mean classification rate), whereas for the combined channel case, the best results were obtained using SVM (83.33% mean classification rate), for the hardest case of differentiating six basic discrete emotions.

In Ref. [19], *IF-THEN* rules of a neurofuzzy system detecting positive and negative emotions are discussed. The study presents the individual performance (ranging from 60 to 82%) of the system for the recognition of emotions (two or four categories) of 11 participants. The decision process is organized into levels where fuzzy membership functions are calculated and combined to achieve decisions about emotional states. The inputs of the system are not only EEG-based features, but also visual features computed on the presented stimulus image.

2.2. Event-related potentials and emotion

Studies of event-related potentials (ERPs) deal with signals that can be tackled at different levels of analysis: signals from single-trials, ensemble averaged signals where the ensemble encompasses several single-trials and signals resulting from a grand-average over different trials as well as subjects. The segments of the time series containing the single-trial response signals are time-locked with the stimulus: t_i (negative value) before and t_f (positive value) after stimulus onset. The ensemble average, over trials of one subject, eliminates the spontaneous

activity of brain and the spurious noisy contributions, maintaining only the activity that is phase-locked with the stimulus onset. The grand-average is the average, over participants, of ensemble averages and it is used mostly for visualization purposes to illustrate the outcomes of the study. Usually, a large number of epochs linked to the same stimulus type need to be averaged in order to enhance the signal-to-noise ratio (SNR) and to keep the mentioned phase-locked contribution of the ERP. Experimental psychology studies on emotions show that the ERPs have characteristics (amplitude and latency) of the early waves which change according to the nature of the stimuli [20, 21]. In Ref. [16], the characteristics of ensemble-average are the features of the classifier. However, this model can only roughly approximate reality, since it cannot deal with robust dynamical changes that occur in the human brain [22].

Due to the mentioned limitation, frequency analysis is more appropriate, as long as it is assumed that certain events affect specific bands of the ongoing EEG activity. Therefore, several investigations have studied the effect of stimuli on characteristic frequency bands. Hence, these measures reflect changes in gamma (γ), beta (β), alpha (α), theta (θ) or delta (δ) bands and can be used as input to a classification system. It is known that beta waves are connected to an alert state of mind, whereas alpha waves are more dominant in a relaxing context [23]. Alpha waves are also typically linked to expectancy phenomena and it has been suggested that the main sources of them are located in parietal areas, while beta activity is most prominent in the frontal cortex over other areas during intense-focused mental activity [22]. Furthermore, regarding the emotional valence processing, psychophysiological research works have shown different patterns in the electrical activity recorded from the two hemispheres [24]. By comparing the power of spectral bands between the left and the right hemisphere of the brain of one participant, they reveal that the left frontal area is related to positive valence, whereas the right one is more related to negative valence [25].

In brain-related studies, one of the most popular, a simple and reliable measure from the spectral domain is the event-related desynchronization/synchronization (ERD/ERS). It represents a relative decrease (ERD) or increase (ERS) in the power content in time intervals after the stimulus onset when compared to a reference interval defined before the stimulus onset [26]. ERD/ERS estimated for the relevant frequency bands during the perception of emotional stimulus have been analyzed [27, 28]. It is suggested that ERS in the theta band is related to emotional processes, together with an interaction between valence and hemisphere for the anterior-temporal regions [27]. Later on, experiments showed that the degree of emotional impact of the stimulus is significantly associated with increase in evoked synchronization in the δ-, α-, β-, γ- bands [28]. In the same study, it was also suggested that the anterior areas of the cortex of both hemispheres are associated predominantly with the valence dimension of emotion. Moreover, in Ref. [29], it has been suggested that delta and theta bands are involved in distinguishing between emotional and neutral states, either with explicit or implicit emotions. Furthermore, in Ref. [30], the results showed that centrofrontal areas showed significant differences of ERD-delta associated with the valence dimension. They also reported that desynchronization of the medium alpha range is associated with attentional resources. More recently, in Ref. [31], the relationships of the late positive potential (LPP) and alpha-ERD during the viewing of emotional pictures have been investigated. The statistical results obtained by these studies show that it is worth considering ERD/ERS measures as inputs to classifiers

meant to automatically recognize emotions. Interestingly, a recent review about affective computing systems [7] emphasizes the advantages of using frequency-based features instead of the ERP components.

3. Materials and methods

In our valence detection system, we have addressed the problem of selecting the most relevant features to define the scalp region of interest by including a wrapper-based classification block. Feature extraction is based on ERD/ERS measures computed in short intervals and is performed either on signals averaged over an ensemble of trials or on single-trial response signals, in order to carry out inter and intrasubject analysis, respectively. The subsequent wrapper classification stage is implemented using two different classifiers: an ensemble classifier, i.e., a random forest and an SVM. The feature selection of algorithm is wrapped around the classification of algorithm recursively identifying the features which do not contribute to the decision. These features are eliminated from the feature vector. This goal is achieved by applying an importance measure, which depends on the parameters of the classifier. The two variants of the system were implemented in MATLAB also using some facilities of open source software tools like EEGLAB [32], as well as random forest and SVM packages [33].

3.1. Data set

A total of 26 female volunteers participated in the study (age 18-62 years; mean = 24.19; SD = 10.46). Only adult women were chosen in this experiment to avoid gender differences [21, 34, 35]. All participants had normal or corrected to normal vision and none of them had a history of severe medical treatment, neither psychological nor neurological disorders. This study was carried out in compliance with the Helsinki Declaration and its protocol was approved by the Department of Education from the University of Aveiro. All participants signed informed consents before their inclusion.

Each one of the selected participants was comfortably seated at 70 cm from a computer screen (43.2 cm), alone in an enclosed room. The volunteer was instructed verbally to watch some pictures, which appeared on the center of the screen and to stay quiet. No responses were required. The pictures were chosen from the IAPS repository. A total of 24 images with high arousal ratings (>6) were selected, 12 of them with positive affective valence (7.29 ± 0.65) and the other 12 with negative affective valence (1.47 ± 0.24). In order to match as closely as possible the levels of arousal between positive and negative valence stimuli, only high arousal pictures were shown, avoiding neutral pictures. **Figure 1** shows the representation of the stimuli in arousal/valence space.

Three blocks with the same 24 images were presented consecutively and pictures belonging to each block were presented in a pseudorandom order. In each trial, a fixation single cross was presented on the center of the screen during 750 ms, after which an image was presented during 500 ms and finally, a black screen during 2250 ms (total duration = 3500 ms). **Figure 2** shows a scheme of the experimental protocol.
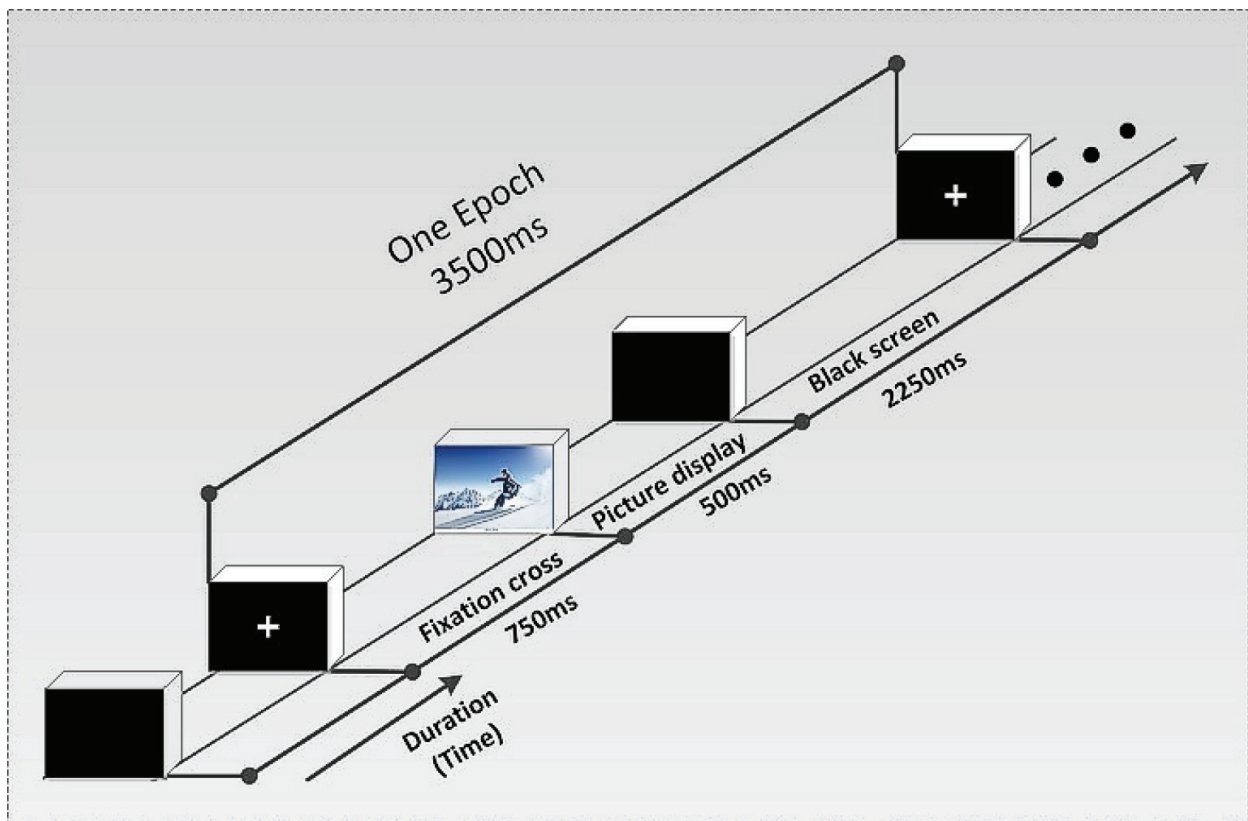
Figure 2. Experimental protocol: series of the stimuli presentation for a complete trial.

EEG activity on the scalp was recorded from 21 Ag/AgCl sintered electrodes (Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, Oz, O_2) mounted on an electrode cap from EasyCap according to the international 10/20 system, internally referenced to an electrode on the tip of the nose. The impedances of all electrodes were kept below 5 kΩ. EEG signals were recorded, sampled at 1 kHz and preprocessed using software Scan 4.3. First, a notch filter centered in 50 Hz was applied to eliminate AC contribution. EEG signals were then filtered using high-pass and low-pass Butterworth filters with cutoff frequencies of 0.1 Hz and 30 Hz, respectively. The signal was baseline corrected and segmented into time-locked epochs using the stimulus onset (picture presentation) as reference. The length of the time windows was 950 ms: from 150 ms before picture onset to 800 ms after it (baseline = 1150 ms).

3.2. Feature extraction

The signals (either single trials or average segments) are filtered by four 4th-order bandpass Butterworth filters. $K = 4$ filters are applied following a zero-phase forward and reverse digital filter methodology not including any transient (see *filtfilt* MATLAB function [36]). The four frequency bands have been defined as: $\delta = Z[0.5, 4]$ Hz, $\theta = Z[4, 7]$ Hz, $\alpha = Z[8, 12]$ Hz and $\beta = Z[13, 30]$ Hz. From a technical point of view, ERD/ERS computation reduces significantly the initial sample size per trial (800 features corresponding to the time instants) to a much smaller number, optimizing the design of the classifier. For each filtered signal, the ERD/ERS is estimated in $I = 9$ intervals following the stimulus onset and with a duration of 150 ms and 50% of overlap between consecutive intervals. The reference interval corresponds to the 150 ms pre-stimulus period. For each interval, the ERD/ERS is defined as

$$f_{ik} = \frac{E_{rk} - E_{ik}}{E_{rk}} = 1 - \frac{E_{ik}}{E_{rk}} \tag{1}$$

where E_{rk} represents the energy within the reference interval and E_{ik} is the energy in the ith interval after stimulus in the kth band, for i = 1,2,…, 9 and k = 1,…,4. Note that when $E_{rk} > E_{ik}$, then f_{ik} is positive, otherwise it is negative. And furthermore notice that the measure has an upper bound $f_{ik} \leq 1$ because energy is always a positive value. Energies E_{ik} are computed by adding up instantaneous energies within each of the I = 9 intervals of 150 ms duration. The energy E_{rk} is estimated in an interval of 150 ms duration defined in the pre-stimulus period. Generally, early poststimulus components are related to an increase of power in all bands due to the evoked potential contribution and this increase is followed by a general decrease (ERD), especially in the alpha band, which can be modulated by a perceptual enhancement as a reaction to relevant contents by the presence of high arousal images [31].

In summary, each valence condition can be characterized by f_{ikc}, where i is the time interval, k is the characteristic frequency band and c refers to the channel. A total of M = $I \times K \times C$ = 9 × 4 × 21 = 756 features is computed for the multichannel segments related to one condition. Following, the features f_{ikc} will be concatenated into a feature vector with components f_m, m = 1,…,M, with M = 756.

3.3. Classification using wrapper approaches

The target of any feature selection method is the selection of the most pertinent feature subset which provides the most discriminant information from a complete feature set. In the wrapper approach, the feature selection algorithm acts as a wrapper around the classification algorithm. In this case, the feature selection consists of searching a relevant subset of features from high-dimensional data sets using the induction algorithm itself as part of the function-evaluating features [37]. Hence, the parameters of the classifier serve as scores to select (or to eliminate) features and the corresponding classification performance is the guide to an iterative procedure. The recursive feature elimination strategy using a linear SVM-based classifier is a wrapper method usually called support vector machine recursive feature elimination (SVM-RFE) [38]. This strategy was introduced when the data sets had a large number of features compared to the number of training examples [38], but it was recently applied for class-imbalanced data sets [39]. A similar strategy can be applied with other learning algorithms, for instance random-forest that has an embedded method of feature selection. The random forest is an ensemble of binary decision trees where the training is achieved by randomly selecting subsets of features. Therefore, computing a variable using parameters of the classifier, which somehow reflect the importance of each input (feature) of the classifier, an iterative procedure can be developed. Assuming that this variable importance is r_m, the steps of the wrapper method are:

1. Initialize: create a set of indices M = {1,2,…,M} relative to the available features and set $F = M$.

2. Organize data set X by forming the feature vectors with the feature values whose index is in set M, labeling each feature vector according to the class it belongs (negative or positive valence).

3. Compute the accuracy of the classifier using the leave-one-out (LOO) cross-validation strategy.

4. Compute the global model of the classifier using the complete data set X.

5. Compute r_m of the feature set and eliminate from set M the indices relative to the twenty least relevant features.

6. Update the number of features accordingly, i.e. $F \leftarrow F$ -20.

7. Repeat steps 2–6 while the number of features in set M is larger than $M_{\min}$ = 36.

Accuracy is the proportion of true results (either positive or negative valence) in the test set. The leave-one-out strategy assumes that only one example of the data set forms the test set while all the remaining belong to the training set. This training and test procedure is repeated so that all the elements of the data set are used once as test set (step 3 of the wrapper method). Then, after computing the model of the classifier with the complete data, the importance of each feature is estimated (steps 4 and 5).

As mentioned before, random forest and linear SVM are classifiers that can be applied in a wrapper method approach and used to estimate r_m. For convenience, the next two subsections review the relevant parameters of both classifiers and their relation to the variable importance mechanism.

3.3.1. Random forest

The random forest algorithm, developed by Breiman [40], is a set of binary decision trees, each performing a classification, being the final decision taken by majority voting. Each tree is grown using a bootstrap sample from the original data set and each node of the tree randomly selects a small subset of features for a split. An optimal split separates the set of samples of the node into two more homogeneous (pure) subgroups with respect to the class of its elements. A measure for the impurity level is the Gini index. By considering that ω_c, c = 1…C are the labels given to the classes, the Gini index of node i is defined as

$$G(i) = 1 - \sum_{c=1}^{C} \left(P(\omega_c)\right)^2 \tag{2}$$

where $P(\omega_c)$ is the probability of class ω_c in the set of examples that belong to node i. Note that $G(i) = 0$ when node i is pure, e.g., if its data set contains only examples of one class. To perform a split, one feature f_m is tested $f_m > f_0$ on the set of samples with n elements which is then divided into two groups (left and right) with n_l and n_r elements. The change in impurity is computed as

$$\Delta G(i) = G(i) - \left(\frac{n_l}{n} G(i_l) - \frac{n_r}{n} G(i_r)\right) \tag{3}$$

The feature and value that results in the largest decrease of the Gini index is chosen to perform the split at node i. Each tree is grown independently using random feature selection to decide the splitting test of the node and no pruning is done on the grown trees. The main steps of this algorithm are

1. Given a data set T with N examples, each with F features, select the number T of trees, the dimension of the subset $L < F$ of features and the parameter that controls the size of the tree (it can be the maximum depth of the tree, the minimum size of the subset in a node to perform a split).

2. Construct the $t = 1 \ldots T$ trees.

    a. Create a training set $T\,t$ with N examples by sampling with replacement the original data set. The out-of-bag data set Ot is formed with the remaining examples of T not belonging to $T\,t$.

    b. Perform the split of node i by testing one of the $L = \sqrt{F}$ randomly selected features.

    c. Repeat step 2b until the tree t is complete. All nodes are terminal nodes (leafs) if the number n_s of examples is $n_s \leq 0.1N$.

3. Repeat step 2 to grow next tree if $t \neq T$. In this work $T = 500$ decision trees were employed.

After training, the importance r_m of each feature f_m in the ensemble of trees can be computed by adding the values of $\Delta G(i)$ of all nodes i where the feature f_m is used to perform a split. Sorting the values r by decreasing order, it is possible to identify the relative importance of the features. The $F = 20$ least relevant features are eliminated from the feature vector f.

3.3.2. Linear SVM

Linear SVM parameters define decision hyperplanes or hypersurfaces in the multidimensional feature space [41, 42], that is:

$$g(w) = w^T x + b = 0 \tag{4}$$

where $x \equiv f$ denotes the vector of features, w is known as the weight vector and b is the threshold.

The optimization task consists of finding the unknown parameters w_m, $m = 1,\ldots,F$ and b [43]. The position of the decision hyperplane is determined by vector w and b: the vector is orthogonal to the decision plane and b determines its distance to the origin. For the Linear SVM the vector w can be explicitly computed and this constitutes an advantage as it decreases the complexity during the test phase. With the optimization algorithm the Lagrangian values, $0 \leq \lambda_i \leq C$ are estimated [43]. The training examples, known as support vectors, are related with the nonzero Lagrangian coefficients. The weight vector then can be computed

$$w = \sum_i^{N_s} y_i \lambda_i x_i \tag{5}$$

where N_s is the number of the support vectors and (x_i, y_i) is the support vector and corresponding label $\{-1,1\}$. The threshold b is estimated as an average of the projected supported vectors $w^T x_i$ corresponding to $C \neq 0$. The value of C needs to be assigned to run the training optimization algorithm and controls the number of errors allowed versus the margin width. During the optimization process, C represents the weight of the penalty term of the optimization function that is related to the misclassification error in the training set. There is no optimal procedure to assign this parameter but it has to be expected that:

– If C is large, the misclassification errors are relevant during optimization. A narrow margin has to be expected.

– If C is small, the misclassification errors are not relevant during optimization. A large margin has to be expected.

Note that this is important because in a real application linearly separable problems are not to be expected and it is more realistic to perform an optimization where misclassifications are allowed. In the following simulations, the parameter $C = 1$ and the software MATLAB is used [44].

The relevance of the mth entry of the feature vector is then determined by the corresponding value w_m in the weight vector. In particular if $\left| w_m \right| \neq 0$, the corresponding feature do not contribute to the value of $g(w)$ [38]. Then, setting $r_m \equiv w_m$ for the SVM classifier and sorting the absolute values, the importance of the features is found out.

4. Results and discussion

To ease interpreting the following results, it is possible to link wrapper methods to some statistical contrasts (e.g. t-test) used by psychologists to test which EEG features change depending on the experimental condition. Note that in the two cases the goal is to perform a dimension reduction before the classification step. For instance, another alternative methodology would consist of transforming the initial vector of features to low-dimensional space by performing a singular value decomposition [45]. Both approaches can be considered as filter techniques to reduce the dimension of an initial feature vector. In the former, statistical analysis, the dimension reduction is achieved according to a parameter from a classifier and each feature is taken individually to check how its value influences the classification outcome. In the latter, machine learning approach, the significant features, selected from the initial vector, are obtained after comparing two sets of features belonging to two different conditions and checking a statistical value. Classification techniques have the advantage of dealing with the set of features as a whole without needing a complementary observation (belonging to another condition). Therefore, the results obtained by wrapper methods can complement the conclusions drawn by applying other statistical tests, indicating the most relevant features related to specific processes, e.g., affective valence processing.

Considering feature elimination and the concomitant number of relevant features, as can be seen from **Figures 3–6**, the accuracy of the wrapper classifiers improves with a decreasing number of relevant features in both, inter or intrasubject classification strategies. In all cases, the system achieves 80% accuracy rate using random forest whereas the system reaches values close to 100% by means of SVM when the classifier has less than 100 relevant features as input.

4.1. Intersubject classification

Figures 3 and **4** show the accuracy versus the number of removed features obtained by applying the two methods: random forest and SVM, respectively. The accuracy was computed with a LOO cross-validation strategy and a total of 52 feature vectors were involved, which represent the ensemble averages referring to positive and negative affective valence responses of all volunteers investigated (26 for each class). Each feature vector is composed of $M = 756$ elements (see section 3.2). A global accuracy of 79% is achieved by the system if roughly 500 irrelevant features are removed from the input feature set with random forest, whereas the system yields an accuracy peak value of 100% using SVM-RFE after removing 680 features, remaining less than 100 features as relevant ones.
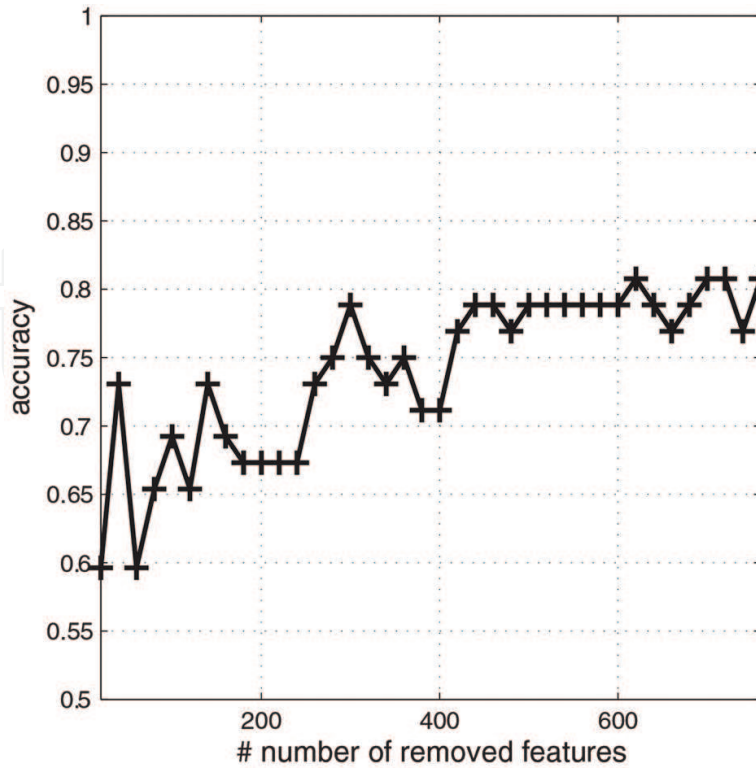
Figure 3. Intersubject accuracy obtained by the implemented random forest. Features extracted from ensemble-average signals are computed at least over 30 single trials.
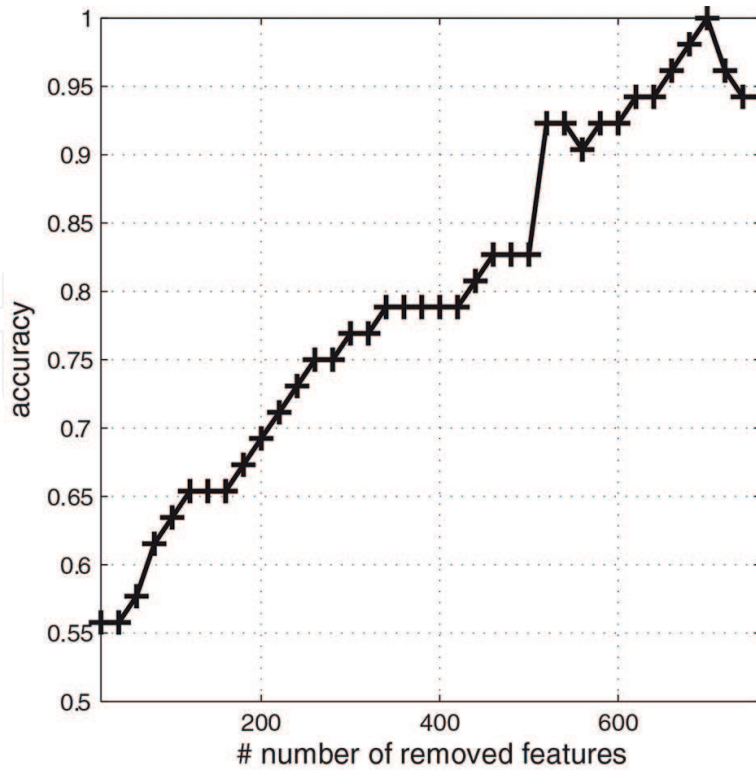

Figure 4. Intersubject accuracy obtained by the implemented SVM-RFE. Features extracted from ensemble-average signals are computed at least over 30 single trials.
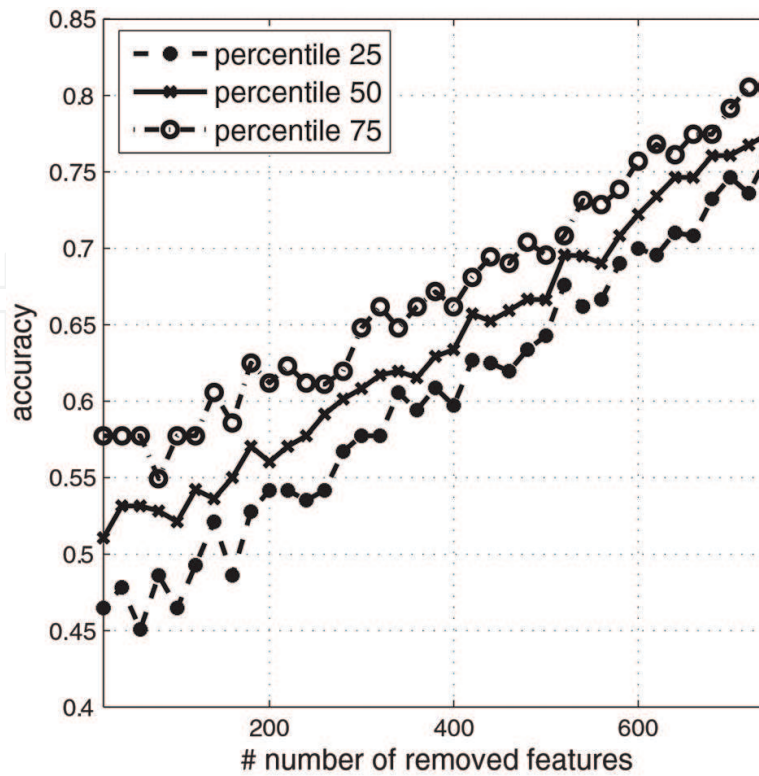
Figure 5. Intrasubject percentiles of accuracy values versus the number of features removed using random forest. The last point corresponds to 36 selected features.
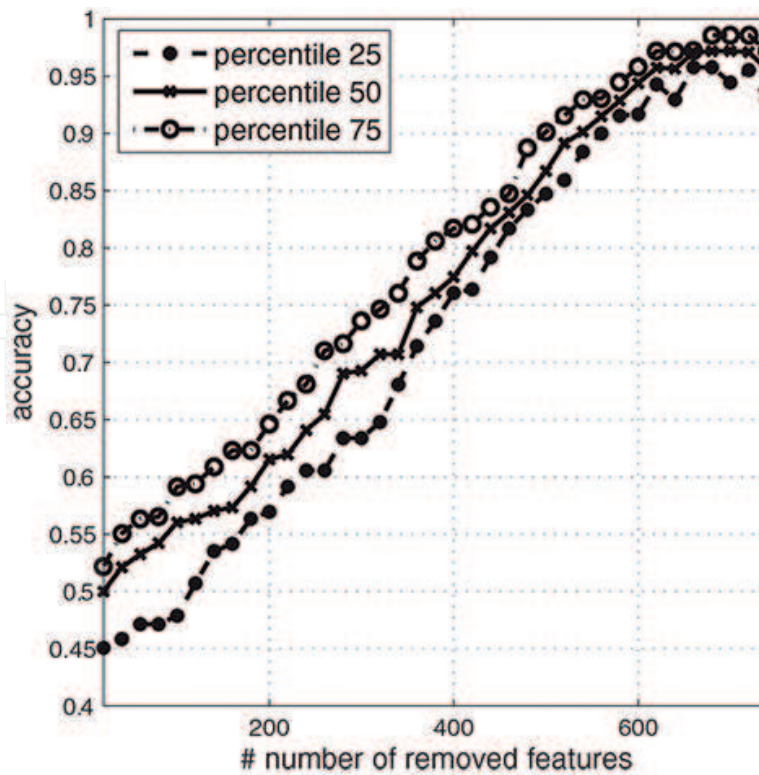

Figure 6. Intrasubject percentiles of accuracy values versus the number of features removed using SVM-RFE. The last point corresponds to 36 selected features.

Tables 1 and **2** describe the spatial and temporal locations of the relevant features when the input of the classifiers is the data set formed by these 52 feature vectors. Concerning spatial locations, both random forest and the SVM algorithm allocate the relevant features consistently in frontal regions of the brain, although SVM also keeps a significant number from centroparietal regions. This corroborates other research works where, during affective processing, the particular contribution of frontal regions has also been pointed out [46, 47]. Concerning location in time, with a random forest most of the features display *medium* and *long* latencies while with an SVM the most relevant time interval corresponds to *medium* latencies. Hence, in contrast to a random forest, the SVM selects a larger number of features from early poststimulus time intervals. These results also match previous brain studies reported in literature, in which ensembles of averaged signals were used as well [20, 48]. Note that although the two methods hardly agree to the time intervals where features show up, both highlight frontal areas as relevant spatial locations for affective valence processing.

Scalp region	Beta	Alpha	Theta	Delta	Total
Frontal	7	2	4	5	18
Central-temporal	6	0	3	0	9
Parietooccipital	5	2	0	2	9
Time interval	Beta	Alpha	Theta	Delta	Total
Short	0	1	0	0	1
Medium	6	1	0	2	9
Long I	12	0	1	3	16
Long II	0	2	6	2	10

Upper table: Space location (EEG channels): frontal (FP1, FPz, FP2, F7, F3, Fz, F4, F8), central-temporal (T7, C3, Cz, C4, T8) and parietal-occipital (P7, P3, Pz, P4, P8, O1, Oz, O_2). Lower table: Time location (time intervals): short ($i = \{1,2\}$), medium ($i = \{3.4\}$), long l ($i = \{5,6\}$) and long II ($i = \{7,8,9\}$).

Table 1. Distribution of the 36 selected features within each band by random forest (intersubject classification).

4.2. Intrasubject classification

Figures 5 and **6** show the global accuracy, computed by averaging the particular accuracy values of all participants, when the classifiers were trained on only one subject's data. For an intrasubject classification purpose, features were extracted from single-trial signals as described above (Section 3). The training set for each subject is made up by a total of 65-72 single trials for both classes of emotions and LOO cross-validation strategy is applied as well.

Similar to intersubject analysis, SVM-RFE yielded better results in terms of accuracy rates than random forest when features are extracted from single-trials. SVM-RFE reaches mean values close to the maximal accuracy and up to 100% for some subjects. Nevertheless, random forest keeps an 80% accuracy as the upper limit.

Scalp region	Beta	Alpha	Theta	Delta	Total
Frontal	5	6	0	5	16
Central-temporal	6	0	2	5	13
Parietooccipital	3	3	1	0	7
Time interval	**Beta**	**Alpha**	**Theta**	**Delta**	**Total**
Short	2	3	3	1	9
Medium	6	2	0	3	11
Long I	4	0	0	3	7
Long II	2	4	0	3	9

Upper table: Space location (EEG channels): frontal (Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8), central-temporal (T7, C3, Cz, C4, T8) and parietal-occipital (P7, P3, Pz, P4, P8, O1, Oz, O2). Lower table: Time location (time intervals): short ($i = \{1,2\}$), medium ($i = \{3.4\}$), long l ($i = \{5,6\}$) and long II ($i = \{7,8,9\}$).

Table 2. Distribution of the 36 selected features within each band by SVM-RFE (intersubject classification).

A comparison of the outcomes of individual training sessions, with respect to the 36 features that remain, reveals a large interindividual. All training sessions encompassed an equal number of iterations. For each feature, it was then counted how often it occurred in any subject. **Figure 7** displays this comparison. It shows that, for example, 220 features never survived
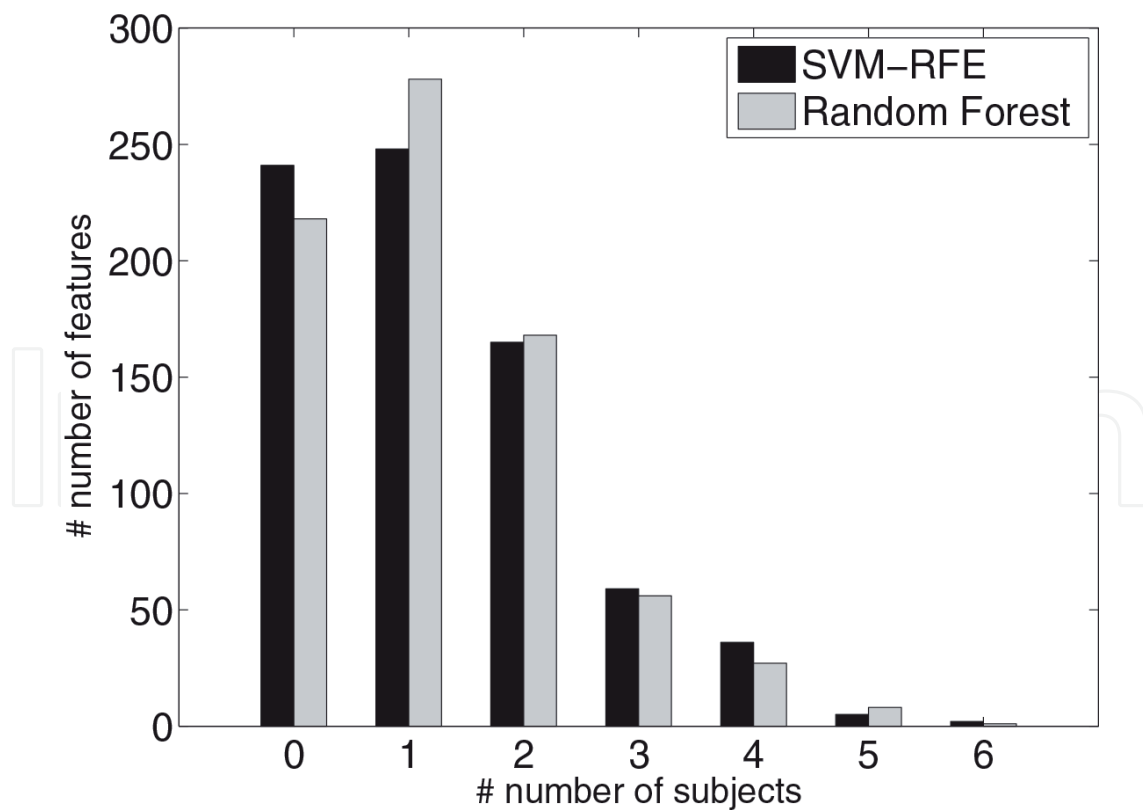

Figure 7. Within the 36 features selected from each individual training, the histogram counts the number of times a feature was selected using both wrapper methods.

in any individual thus may be considered completely irrelevant. Remarkably, few features appear consistently as relevant features in at least six out of 26 subjects confirming the high interindividual heterogeneity, independently on the applied method for selecting features. A similar conclusion was drawn in Ref. [15], in this case by using a feature selection block before performing classification. However, note that a comparable accuracy value is achieved whether decision making is based on a set of 52 feature vectors (ensemble averages over trials and subjects) or on training classifiers individually with 65–72 feature vectors for each subject.
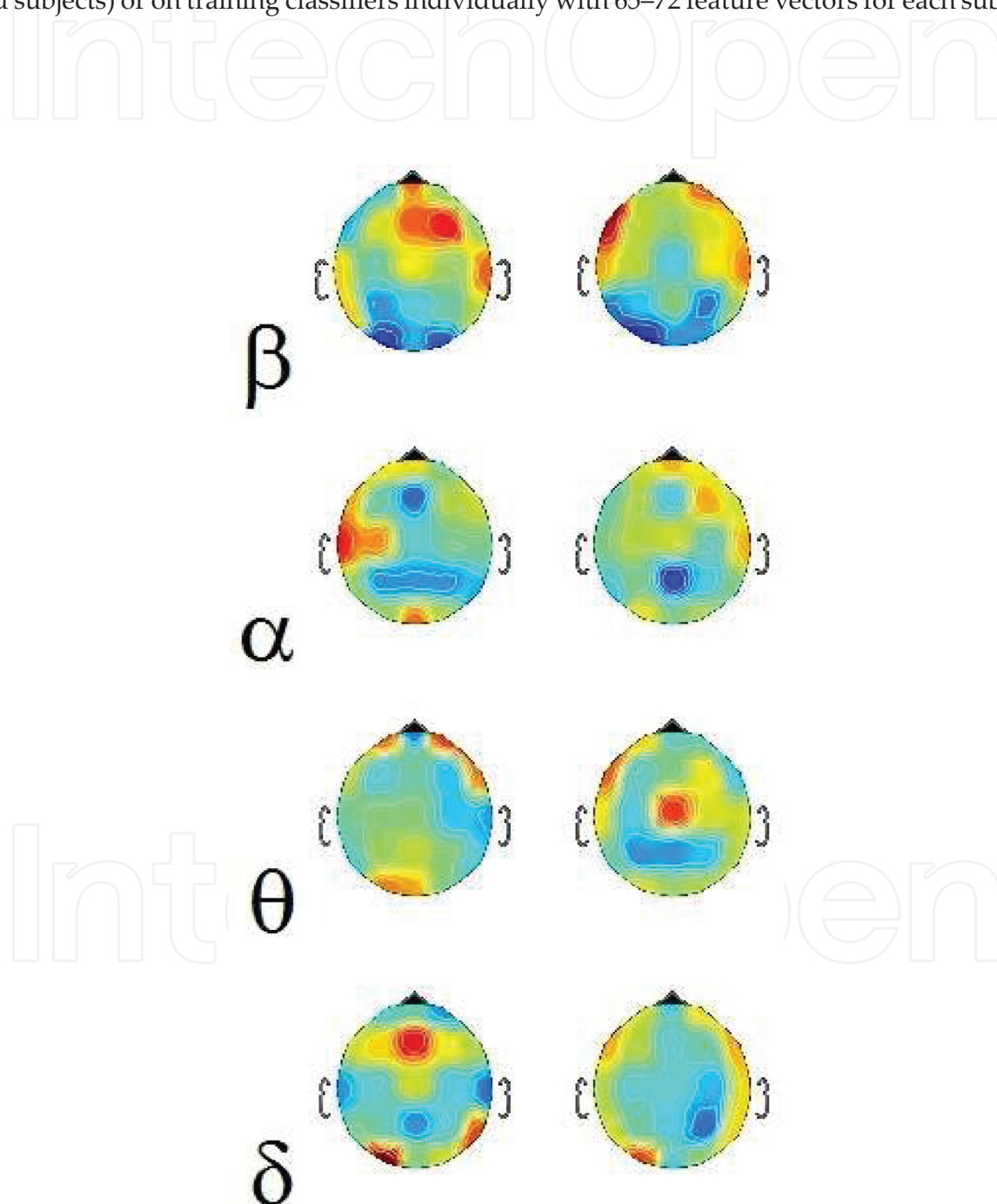

Figure 8. Spatial location of feature relevance in each frequency band obtained from counting the contribution from all subjects within intrasubject classifications (left column: random forest, right column: SVM). The relevance is represented by a color map, where blue represents the least relevant features (nonselected features) and red represents the most relevant ones (selected as relevant by all subjects).
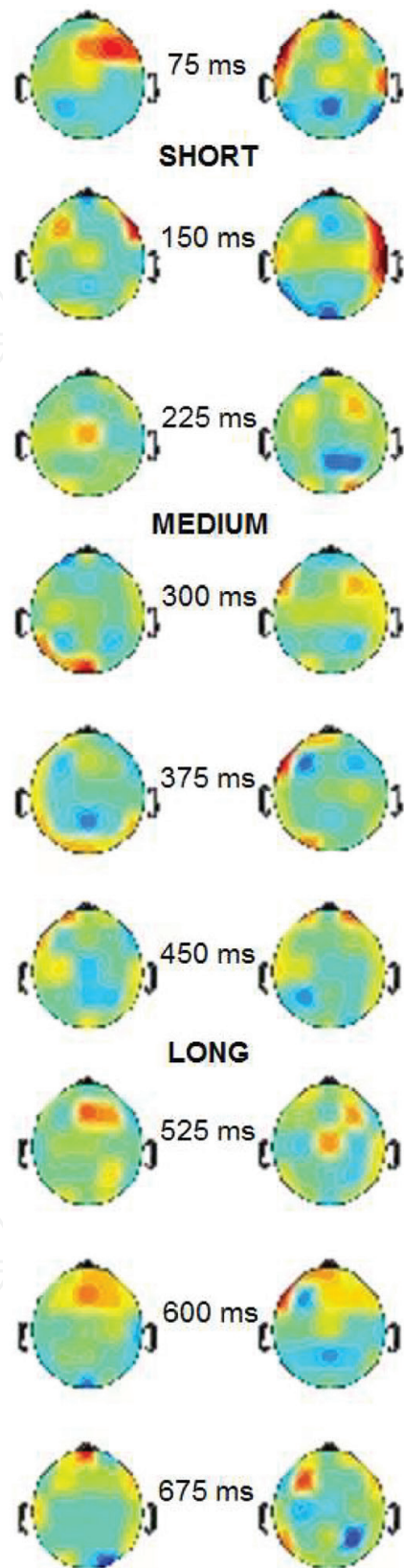
Figure 9. Relevance of the features selected for different latencies and spatial locations (left column: random forest, right column: SVM) following from counting the contribution of all subjects within an intrasubject classification. Feature importance is visualized by a normalized color map, where blue represents the least relevant features (nonselected features) and red represents the most relevant ones (selected as relevant by all subjects).

Figures 8 and **9** show the relevance of the features chosen by the two methods in topographical maps. As can be seen from **Figure 8**, on average, both algorithms allocate the most relevant features in the frontal region in agreement with intersubject applications. Similarly, both also identify relevant features mostly in the beta bands. According to **Figure 9**, both algorithms allocate important features showing up with short latencies in the frontal areas of the brain. Concerning *medium* and *long* latencies both algorithms again identify important features in frontal areas though their importance is more pronounced with the random forest.

Although intersubject and intrasubject methodologies show a similar performance, they have different application scenarios. The intersubject classification is mostly suitable for offline applications as well as for brain studies in order to complement the statistical methods. For instance, in Ref. [49], an SVM-RFE scheme was exclusively applied to identify scalp spectral dynamics linked with the affective valence processing and to compare with standard statistical results (*t*-test). In that work, a different technique for feature extraction was developed, whose goals consisted of creating a particular volume of features by means of a wavelet filtering. In this way, a high-dimensional data set was represented by means of three dimensions: frequency (resolution: 1 Hz), time (resolution: 1 ms) and topographical location (21 EEG channels).

Due to the biological variability observed, intrasubject studies cannot generalize easily across a cohort of subjects. Thus, intrasubject approaches might be interesting for personalized studies where subjects need to be followed up for a couple of sessions, such as in a rehabilitation therapy, or for neurofeedback-based applications. An example of an intrasubject study is reported in Ref. [50], where the neuroticism trait is analyzed using EEG to check the influence of individual differences in the emotional processing and the susceptibility of each brain region. In that work SVM was used as well, although from a different standpoint, since it was performed in subject identification tasks from single trials.

5. Conclusions

A novel valence recognition system has been presented and applied to EEG signals, which were recorded from volunteers subjected to emotional states elicited by pictures drawn from IAPS repository. A cohort of 26 female participants has been investigated. The recognition system encompasses a feature extraction stage and a classification module including feature elimination. The complete system focused on both an intersubject and an intrasubject situation. Both studies show a similar performance with regard to the classification accuracy. The recursive feature elimination (selection) was designed based on a random forest classifier or support vector machine and increased the initial classification accuracy in a range from 20% to 45%. The importance measures from both algorithms point to frontal areas although no consistent set of features and related latencies could be identified.

This fact points toward a large biological variability of the set of relevant features corresponding to the valence of the emotional states involved. In any case, the classification accuracy achieved compares well with or is even superior to competing systems reported in the literature.

Comparing both classifiers, the SVM achieves better classification accuracy, yielding up to 100% accuracy rate for several subjects by using the selected features for intrasubject classification task, outperforming random forest that reached an 81% maximum peak accuracy. Furthermore, the presented wrapper methods are a good option to greatly reduce the dimension of the input space and deserve to be considered as an alternative for discriminating relevant scalp regions, frequency bands and time intervals from EEG recordings.

Author details

Antonio R. Hidalgo-Muñoz[1]*, Míriam M. López[1], Isabel M. Santos[2], Manuel Vázquez-Marrufo[3], Elmar W. Lang[4] and Ana M. Tomé[1]

*Address all correspondence to: arhidalgom@gmail.com

1 IEETA, University of Aveiro, Aveiro, Portugal

2 CITENSIS, Department of Education and Psychology, University of Aveiro, Aveiro, Portugal

3 University of Seville, Seville, Spain

4 Biophysics CIML Group, University of Regensburg, Regensburg, Germany

References

[1] S. A. Calvo, S. D'Mello. Affect detection: an interdisciplinary review of models, methods and their applications. IEEE Transactions on Affective Computing. 2010;**1**(1):18–37.

[2] J. L. Salmerón. Fuzzy cognitive maps for artificial emotions forecasting. Applied Soft Computing. 2012;**12**:3704–3710.

[3] M. Krauledat, K. Grzeska, M. Sagebaum, B. Blankertz, C. Vidaurre, K. R. Miller, M. Schrïder. Playing pinball with non-invasive BCI. In: Advances in Neural Information Processing Systems 21; 2009. pp. 1641–1648. Edited by D. Koller and D. Schuurmans and Y. Bengio and L. Bottou Publisher: Curran Associates, Inc. New York (USA).

[4] D. Huang, K. Qian, D. Fei, W. Jia, X. Chen, O. Bai. Electroencephalography (EEG)-based brain-computer interface (BCI): A 2-D virtual wheelchair control based on event-related desynchronization/synchronization and state control. IEEE transactions on Neural Systems and Rehabilitation engineering. 2012;**20**(3):379–388.

[5] J. A. Russell. A circumplex model of affect. Journal of Personality and Social Psychology. 1980;**39**(6):1161–1178.

[6] M. M. Bradley, P. J. Lang. The international affective picture system (IAPS) in the study of emotion and attention. In: M. M. Bradley, P. J. Lang, editors. Handbook of Emotion Elicitation and Assessment. Oxford University Press; Oxford. 2007. pp. 29–46.

[7]   C. Mühl, B. Allison, A. Nijholt, G. Chanel. A survey of affective brain computer interfaces: principles, state-of-the-art and challenges. Brain-Computer Interfaces. 2014;**1**(2):62–84.

[8]   J. Klein, Y. Moon, R. W. Picard. This computer responds to user frustration: theory, design and results. Interacting with Computers. 2002;**14**(2):119–140.

[9]   P. Lang, M. Bradley, B. Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report. 2008.

[10] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras. DEAP: a database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing. 2012;**3**(1):18–31.

[11] R. W. Picard, E. Vyzas, J. Healey. Toward machine emotional intelligence: analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001;**23**(10):1175–1191.

[12] A. Haag, S. Goronzy, P. Schaich, J. Williams. Emotion recognition using bio-sensors: First steps towards an automatic system. In: E. André, L. Dybkjær, W. Minker, P. Heisterkamp, editors. Affective Dialogue Systems. 6th ed. Springer Berlin Heidelberg; Berlin. 2004. pp. 36–48.

[13] K. H. Kim, S. W. Bang, S. R. Kim. Emotion recognition system using short-term monitoring of physiological signals. Medical & Biological Engineering & Computing. 2004;**42**(3):419–427.

[14] K. Schaaff, T. Schultz. Towards emotion recognition from electroencephalographic signals. In: IEEE, editor. 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII), 2009; 2009. pp. 1–6. (sede física de IEEE: New York).

[15] M. Macas, M. Vavrecka, V. Gerla, L. Lhotska. Classification of the emotional states based on the EEG signal processing. In: IEEE, editor. 9th International Conference on Information Technology and Applications in Biomedicine; November 2009; 2009. pp. 1–4. New York.

[16] 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. On the classification of emotional biosignals evoked while viewing affective pictures: An integrated data-mining-based approach for healthcare applications. IEEE Transactions on Information Technology in Biomedicine. 2010;**14**(2):309–318.

[17] P. Petrantonakis, L. Hadjileontiadis. Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. IEEE Transactions on Affective Computing. 2010;**1**(2):81–97.

[18] P. C. Petrantonakis, L. J. Hadjileontiadis. Emotion recognition from EEG using higher order crossings. IEEE Transactions on Information Technology in Biomedicine. 2010;**14**(2):186–197.

[19] Q. Zhang, M. Lee. Emotion development system by interacting with human EEG and natural scene understanding. Cognitive Systems Research. 2012;**14**(1):37–49.

[20] J. K. Olofsson, S. Nordin, H. Sequeira, J. Polich. Affective picture processing: an integrative review of ERP findings. Biological Psychology. 2008;**77**(3):247–265.

[21] C. Lithari, C. Frantzidis, C. Papadelis, A. B. Vivas, M. Klados, C. Kourtidou-Papadeli, C. Pappas, A. Ioannides, P. Bamidis. Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. Brain Topography. 2010;**23**(1):27–40.

[22] X. J. Wang. Neurophysiological and computational principles of cortical rhythms in cognition. Physiological Reviews. 2010;**90**(3):1195–1268.

[23] W. J. Ray, H. W. Cole. EEG alpha activity reflects attentional demands and beta activity reflects emotional and cognitive processes. Science. 1985;**228**(4700):750–752.

[24] E. Beraha, J. Eggers, C. H. Attar, S. Gutwinski, F. Schlagenhauf, M. Stoy, P. Sterzer, T. Kienast, A. Heinz, F. Bermpohl. Hemispheric asymmetry for affective stimulus processing in healthy subjects – a fMRI study. PLoS One. 2012;**7**(10):e46931.

[25] H. J. Yoon, S. Y. Chung. EEG spectral analysis in valence and arousal dimensions of emotion. In: 11th International Conference on Control, Automation and Systems (ICCAS); October 2011. pp. 1319–1322 IEEE.

[26] G. Pfurtscheller, F. H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. Clinical Neurophysiology. 1999;**110** (11):1842–1857.

[27] L. Aftanas, A. Varlamow, S. Pavlov, V. Makhnev, N. Reva. Affective picture processing: event-related synchronization within individually defined human theta band is modulated by valence dimension. Neuroscience Letters. 2001;**303**(2):115–118.

[28] L. I. Aftanas, N. V. Reva, A. A. Varlamov, S. V. Pavlov, V. P. Makhnev. Analysis of evoked EEG synchronization and desynchronization in conditions of emotional activation in humans: temporal and topographic characteristics. Neuroscience and Behavioral Physiology. 2004;**34**(8):859–867.

[29] G. Knyazev, J. Slobodskoj-Plusnin, A. Bocharov. Event-related delta and theta synchronization during explicit and implicit emotion processing. Neuroscience. 2009;**164**(4):1588–1600.

[30] M. A. Klados, C. Frantzidis, A. B. Vivas, C. Papadelis, C. Lithari, C. Pappas, P. D. Bamidis. A framework combining delta event-related oscillations (EROs) and synchronisation effects (ERD/ERS) to study emotional processing. In: Computational Intelligence and Neuroscience; The ACM Digital Library. January; 2009. p. 12.

[31] A. De Cesarei, M. Codispoti. Affective modulation of the LPP and alpha-ERD during picture viewing. Psychophysiology. 2011;**48**(10):1397–1404.

[32] A. Delorme, S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods. 2004;**134**(1):9–21.

[33] A. Jaiantilal, http://code.google.com/p/randomforest-matlab/, 2010.

[34] A. H. Kemp, R. B. Silberstein, S. M. Armstrong, P. J. Nathan. Gender differences in the cortical electrophysiological processing of visual emotional stimuli. Neuroimage. 2004; **21**(2):632–646.

[35] K. Schaaff. Challenges on emotion induction with the international affective picture system. 2008. [Online]. Available: http://csl.ira. uka.de/fileadmin/media/publication files/ SA-Schaaff.pdf

[36] Mathworks, 2012. [Online]. Available: http://www.mathworks.com/help/signal/ref/filt-filt.html

[37] R. Kohavi and G. H. John. Wrappers for feature subset selection. Artificial Intelligence. 1997;**97**(1–2):273–324.

[38] I. Guyon, J. Weston, S. Barnhill, V. Vapnik. Gene selection for cancer classification using support vector machines. Machine Learning. 2002;**46**(1–3):389–422.

[39] S. Maldonado, R. Weber, F. Famili. Feature selection for high-dimensional class-imbalanced data sets using support vector machines. Information Sciences. 2014;**286**(0):228–246. [Online]. Available: http://www.sciencedirect.com/science/article/ pii/S0020025514007154

[40] L. Breiman. Random forests. Machine Learning. 2001;**45**(2): 5–32.

[41] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery. 1998;**2**(2):121–167.

[42] I. A. Illan, J. M. Gorriz, M. M. Lopez, J. Ramirez, D. Salas-Gonzalez, F. Segovia, R. Chaves, C. G. Puntonet. Computer aided diagnosis of Alzheimer's disease using component based SVM. Applied Soft Computing. 2011;**11**(2):2376–2382.

[43] A. Ben-Hur, C. Ong, S. Sonnenburg, B. Schölkopf, G. Rätsch. Support vector machines and kernels for computational biology. PLoS Computational Biology. 2008;**10**(4):e1000173.

[44] Bioinformatics-Toolbox, 2012. [Online]. Available: http://www. mathworks.com

[45] S. N. Daimi, G. Saha. Classification of emotions induced by music videos and correlation with participants rating. Expert Systems with Applications. 2014;**41**(13):6057–6065. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0957417414001882

[46] R. J. Davidson, W. Irwin. The functional neuroanatomy of emotion and affective style. Trends in Cognitive Sciences. 1999;**3**(1):11–21.

[47] S. K. Sutton, R. J. Davidson. Prefrontal brain electrical asymmetry predicts the evaluation of affective stimuli. Neuropsychologia. 2000;**38**(13):1723–1733.

[48] L. R. R. Gianotti, P. L. Faber, M. Schuler, R. D. Pascual-Marqui, K. Kochi, D. Lehmann. First valence, then arousal: the temporal dynamics of brain electric activity evoked by emotional stimuli. Brain Topography. 2008;**20**(3):143–156.

[49] A. R. Hidalgo-Muñoz, M. M. Lopez, I. M. Santos, A. T. Pereira, M. Vazquez-Marrufo, A. Galvao-Carmona, A. M. Tome. Application of SVM-RFE on EEG signals for detecting the most relevant scalp regions linked to affective valence processing. Expert Systems with Applications. 2013;**40**(6):2102–2108.

[50] A. R. Hidalgo-Muñoz, A. T. Pereira, M. M. Lopez, A. Galvao-Carmona, A. M. Tome, M. Vazquez-Marrufo, I. M. Santos. Individual EEG differences in affective valence processing in women with low and high neuroticism. Clinical Neurophysiology. 2013;**124**(9):1798–1806.