EEG study on affective valence elicited by novel and familiar pictures using ERD/ERS and SVM-RFE

Article in Medical & Biological Engineering & Computing · November 2013

DOI: 10.1007/s11517-013-1126-6 · Source: PubMed

Med Biol Eng Comput
DOI 10.1007/s11517-013-1126-6

ORIGINAL ARTICLE

EEG study on affective valence elicited by novel and familiar pictures using ERD/ERS and SVM-RFE

A. R. Hidalgo-Muñoz · M. M. López · A. Galvao-Carmona · A. T. Pereira · I. M. Santos · M. Vázquez-Marrufo · A. M. Tomé

Received: 14 March 2013 / Accepted: 8 November 2013

Abstract EEG signals have been widely explored in emotional processing analyses, both in time and frequency domains. However, in such studies, habituation phenomenon is barely considered in the discrimination of different emotional responses. In this work, spectral features of the event-related potentials (ERPs) are studied by means of event-related desynchronization/synchronization computation. In order to determine the most relevant ERP features for distinguishing how positive and negative affective valences are processed within the brain, support vector machine-recursive feature elimination is employed. The proposed approach was applied for investigating in which way the familiarity of stimuli affects the affective valence processing as well as which frequency bands and scalp regions are more involved in this process. In a group composed of young adult women, results prove that parietooccipital region and theta band are especially involved in the processing of novelty in emotional stimuli. Furthermore, the proposed method has shown to perform successfully using a moderated number of trials.

Keywords Affective valence · EEG · ERD/ERS · Habituation · SVM-RFE

Abbreviations

BCI	Brain–computer interface
ERPs	Event-related potentials
ERD/ERS	Event-related desynchronization/ synchronization
ANOVA	Analysis of variance
SVM-RFE	Support vector machine-recursive feature elimination
IAPS	International affective picture system
NNC	Novel negative condition
NPC	Novel positive condition
FNC	Familiar negative condition
FPC	Familiar positive condition
LR	Left–right
FP	Frontal–parietooccipital
LOO	Leave one out

A. R. Hidalgo-Muñoz (✉) · A. Galvao-Carmona · M. Vázquez-Marrufo
Department of Experimental Psychology, University of Seville, 41018 Seville, Spain
e-mail: arhidalgom@gmail.com

M. M. López
IEETA, University of Aveiro, 3810-193 Aveiro, Portugal

A. T. Pereira · I. M. Santos
Department of Education, University of Aveiro, 3810-193 Aveiro, Portugal

A. T. Pereira · I. M. Santos
IBILI-Institute of Biomedical Research in Light and Image, Faculty of Medicine, University of Coimbra, Coimbra, Portugal

A. M. Tomé
DETI/IEETA, University of Aveiro, 3810-193 Aveiro, Portugal

1 Introduction

Emotion recognition is an important issue, trendy topic in a large number of research works that deal with psychiatric evaluation and applications such as brain–computer interface (BCI), among others [7, 29]. The dimensional model of emotions asserts that emotions can be defined by two dimensions: arousal and affective valence. Arousal is related to the alertness level induced by a stimulus, whereas affective valence is related to the pleasure or displeasure produced by the presentation of the stimulus. The latter is an interesting, essential dimension in studies dealing with emotion processing [18, 30], and is influenced by several variables. One of such variables is habituation, which consists of a reduction in the response to a stimulus when it is repeatedly presented [2, 21].

In [28], novelty is reported as one important dimension in the categorization of emotional stimuli. In the study of habituation phenomenon, a special attention has been paid to electrophysiological signals by the researchers in the field [3, 5, 6, 22]. The majority of the works that use EEG signals to study habituation are focused in analyzing the amplitude and latency of significant peaks of event-related potentials (ERPs) in time domain, e.g., P2, N2, and P3 components [19]. Nonetheless, the huge variability among subjects and the high number of trials needed for the analysis often become an important drawback.

Frequency analysis is an alternative to obtain relevant information and has been proposed for many applications [4, 30]. There are several methods to obtain spectral information from EEG signals: wavelet filtering [8, 10], analysis of coherence or temporal spectral evolution computation among others. One of the most popular, simple and reliable measure is the event-related desynchronization/synchronization (ERD/ERS), which consists of a power decrease/increase, respectively, that occurs after the stimulus onset in defined ranges of frequencies [20]. It has been suggested that ERD/ERS in different ranges of frequency (especially in theta and alpha bands) are modulated by emotions and, specifically, by affective valence [4]. For this reason, we propose ERD/ERS measure as representative features of the ERPs for understanding better the underlying relationship between novelty and affective valence associated with a stimulus. Moreover, the use of a set of representative measures instead of the whole EEG signals or ERPs provide a reduced representation of the high-dimensional signals, being more suitable for dealing with the data in an automatic way as well as for reducing computational costs.

Analyses of variance (ANOVAs) are usually performed in order to find the most relevant differences among ERP conditions. Nonetheless, in some cases, it is not possible to obtain statistically significant differences by means of these analyses, mainly due to sample distribution constraints or to conservative corrections that are needed to be applied on the significance level when multiple comparisons are effectuated. In order to overcome this problem, several methods have been proposed to extract valuable information from EEG [24, 25]. An alternative for identifying relevant features is the use of classifier-based methods. Support vector machine-recursive feature elimination (SVM-RFE) algorithm [9] consists of a *wrapper* method than can be used for selecting the most relevant features for a designed classification task. It internally includes the SVM classification technique [23], which has been successfully used for several clinical and psychological applications [15, 26]. In [10], it is shown that better accuracy values are reached when selecting features from EEG signals by means of SVM-RFE instead of t test. In that research work, the partial energies from each frequency band at each time instant are taken as features and their relevance is evaluated by ranking the features according to both the t values obtained by applying a paired t test and SVM-RFE selection. Then, a classification task is performed, reaching up better results using the second selection method.

In summary, a straightforward method based on SVM-RFE algorithm is presented in this study for selecting the most relevant time intervals and frequency bands whose ERD/ERS are highly linked to affective valence processing. As far as the authors are concerned, this combination has never been used before for studying emotional processes. During the analysis, the influence of the habituation phenomenon is taken into consideration by categorizing each stimulus as novel or familiar. Furthermore, the best option for grouping EEG channels for the valence classification task is studied. This approach is advantageous compared to other conservative statistical method (ANOVA) for analyzing ERPs components, since it takes the complete set of features as a whole entity and works properly using a lower number of trials.

2 Methods

2.1 Database description, signal recording, and preprocessing

Twenty-six healthy female participants were included in this study (age 18–62 years; mean = 24.19; SD = 10.46). Only women have been selected, since it has been demonstrated they are more sensitive to emotional events, especially to negative stimuli [14]. 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. A signed informed consent was obtained from each subject before carrying out the experiment. This study was approved in accordance with the Declaration of Helsinki.

Participants were seated at 70 cm from a matte computer screen (1,280 × 1,024 pixels) (17′) with a refresh rate of 60 Hz. They were instructed to simply visualize the pictures in color (1,024 × 768 pixels) (visual stimuli) that appeared on the center of the screen; therefore, neither additional responses nor judgment were required. All the pictures were chosen from the International Affective Picture System (IAPS) [13]. A total of 24 images with high arousal ratings (>6) were selected: twelve of the pictures elicited positive affective valence (7.29 ± 0.65) and 12 negative affective valence (1.47 ± 0.24). Subjects were asked to score all the images. The scoring results obtained by the participants matched up with the scores given by the standardized version, reaching up a correlation of $\rho = 0.972$, arousing all the pictures high alertness level.

Three blocks with the same 24 images were presented consecutively. The picture order in each block was random to avoid expectancy phenomena or sequence effects. In each trial (total duration = 3,500 ms), a fixation single cross was presented on the center of the screen during 750 ms and then one image was presented during 500 ms and finally a black screen during 2,250 ms.

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, O2) 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 (Compumedics Neuroscan, Germany). Firstly, a notch filter centered in 50 Hz was applied to eliminate AC contribution. EEG signals were then filtered using a fourth-order Butterworth passband filter from 0.1 to 30 Hz and segmented into time-locked epochs using the stimulus onset (picture presentation) as reference and baseline-corrected. The length of the time windows was 950 ms: from 150 ms before picture onset to 800 ms after it (baseline = 150 ms). Artifact rejection was performed based on amplitude values, following a visual inspection of all signals and rejecting epochs that presented clear artifacts and a total of 3 % of trials were discarded. Regarding the ocular correction, a semiautomatic method is applied by following the procedure described in software package Scan 4.3 (Compumedics Neuroscan, Germany). The method employs a regression analysis in combination with artifact averaging to produce a reliable model of the VEOG signal to be subtracted from the EEG channels.

2.2 Experimental conditions and channel grouping

For every subject, four average signals were computed, considering for each one only trials belonging to the same condition (positive or negative valence), and separating trials produced by the first and third time one stimulus appears (novel and familiar stimulus, respectively). Thus, averages were computed using between 10 and 12 single trials as follows:

- Novel negative condition (NNC): Average signal across ERPs induced by negative images from the first block of trials.
- Novel positive condition (NPC): Average signal across ERPs induced by positive images from the first block of trials.
- Familiar negative condition (FNC): Average signal across ERPs induced by negative images from the third (last) block of trials.
- Familiar positive condition (FPC): Average signal across ERPs induced by positive images from the third (last) block of trials.

By dealing with averaged signals, only the evoked contributions from the EEG signals were taken, discarding induced electrophysiological responses that were not phase-locked to the stimulus onset and reducing the impact of other noisy or spurious contributions. Moreover, the averages diminish the original set of signals from each subject to only one sample per condition, getting more precise results in the inter subject analysis.

In addition, topographical averages were also considered, that is, not all the channels were included for each average signal but they were grouped according to different distributions around the scalp. The goal was to obtain one global ERP signal for each different brain region instead of taking all EEG channels in an isolated way. Moreover, this channel grouping method allows to gather together a larger number of EEG signals to compute the averaged ERPs and to mitigate the spurious contributions of noisy artifacts. This kind of channel grouping has been carried out in other research works and provided valuable information about different scalp areas [12].

Two strategies were considered to perform the topographical averages, grouping the signals according to left–right hemispheres (called LR grouping) and frontal–parietooccipital (called FP grouping). The LR grouping was achieved with channels Fp1, F7, F3, T7, C3, P7, P3, O1, representing left hemisphere and channels Fp2, F4, F8, C4, T8, P4, P8, O2, representing right hemisphere. The FP grouping was achieved grouping together channels Fp1, Fp2, Fpz, F7, F3, Fz, F4, F8, from the frontal region and P7, P3, Pz, P4, P8, O1, Oz, O2, from the parietooccipital region. Figure 1 shows an example of the grand averages.

2.3 Feature extraction: ERD/ERS

The power in different bands of the EEG constitutes a measure that allows the quantification of the course of psychological processes. As it is well known, EEG signals are often characterized by decreasing (ERD) or increasing (ERS) the synchrony of underlying neuron networks [20].

In this work, the ERD/ERS is computed using the band power method [12, 20] and represents the portion of band

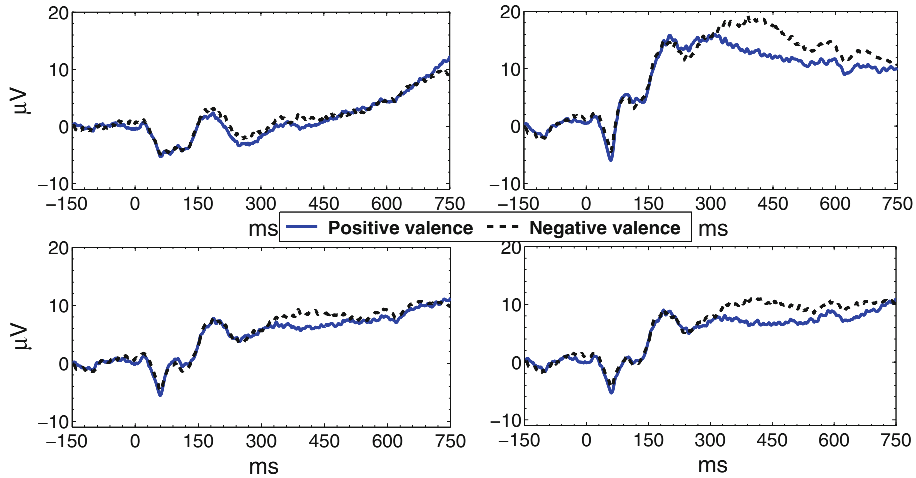

Fig. 1 Grand averages of EEG signals averaged by frontal area (*upper left*), parietooccipital area (*upper right*), left hemisphere (*down left*), and right hemisphere (*down right*) (Task 1: Novel pictures condition)

power that decreases/increases during a test interval Δ_i post-stimulus compared to a reference interval Δ_r pre-stimulus (baseline with the same length) in a specific frequency band. This measure is defined as

$$e_i = (E_{\Delta r} - E_{\Delta i})/E_{\Delta r} \tag{1}$$

where E_{Δ_*} represents the energy of the signal in the corresponding Δ_* time interval. E_{Δ_*} is computed by adding up the square values of voltage levels included in the considered time segment. Note that this relative (to the reference interval) change e_i can be either positive or negative, and as energy values are always positive, the maximum value for e_i is close to 1 when $E_{\Delta i} \approx 0$.

In order to compute the ERD/ERS in different frequency bands, the averaged signals (ERPs) were filtered using a fourth-order zero-phase forward and reverse digital Butterworth passband filter. Band limits were defined as delta band (δ): 0.5–4 Hz, theta band (θ): 4–7 Hz, alpha band (α): 7–12 Hz, beta band (β): 12–30 Hz. Time intervals of 150 ms were taken from the epochs, starting from the image onset (0 ms) and up to 750 ms after image onset, with an overlap of 50 %. A total of 9 time intervals were computed for each epoch, and they were named according to the center instant of the interval (75, 150, 225, 300, 375, 450, 525, 600, and 675). For each band, each time interval corresponds to one ERD/ERS measure.

For visualization purposes, Figs. 2 and 3 show ERD/ERS for each frequency band of the grand-average signal from the frontal and parietooccipital regions, and for each of the four conditions. In these figures, some differences can be observed between novel and familiar conditions in delta and theta bands. Alpha and beta bands are relevant in terms of their relative level of energy, specifically in the time interval from 225 to 450 ms post-stimulus. This fact makes the measure of ERD/ERS be suitable for determining the influence of the habituation and the temporal evolution of the stimulus processing.

2.4 Feature selection: SVM-RFE

SVMs separate a given set of binary labeled training data with a hyperplane that is maximally distant from two classes (ω_1 and ω_2) known as the maximal margin hyperplane. The objective of SVM is to build a function $f : \Re^M \to \{\pm 1\}$ using training data, that is, M-dimensional patterns $\mathbf{x}_i$ and class labels y_i:

$$(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_S, y_S) \in \Re^M \times \{\pm 1\} \tag{2}$$

so that f will correctly classify new examples $(\mathbf{x}, y)$. S is the number of samples in the database. Linear discriminant functions define decision hypersurfaces or hyperplane in a multidimensional feature space, that is,

$$g(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b = 0 \tag{3}$$

where $\mathbf{w}$ is known as the weight vector and b as the threshold. The weight vector $\mathbf{w}$ is orthogonal to the decision hyperplane and b determines the distance of the plane to the origin. The optimization task consists of finding the unknown parameters $w_i, i = 1, \ldots, m$ (entries of weight vector) and b. Given a new element $\mathbf{x}$, the predicted label is then

$$g(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b \Rightarrow \begin{cases} g(\mathbf{x}) > 0, & \mathbf{x} \in \omega_1 \Leftrightarrow \text{label} \equiv 1 \\ g(\mathbf{x}) < 0, & \mathbf{x} \in \omega_2 \Leftrightarrow \text{label} \equiv -1 \end{cases} \tag{4}$$

Note that if there are elements in the vector $\mathbf{w}$ close to zero, they have no influence on $g(\mathbf{x})$ and consequently they do not contribute to the decision. In [9], it was suggested a feature selection technique (SVM-RFE) that eliminates recursively the features corresponding to the lowest values $|w_i|$. The algorithm uses all the features the first time and rejects consecutively the τ features considered less relevant by sorting out the absolute values of the entries w_i. This process is repeated as long as the classifier performance

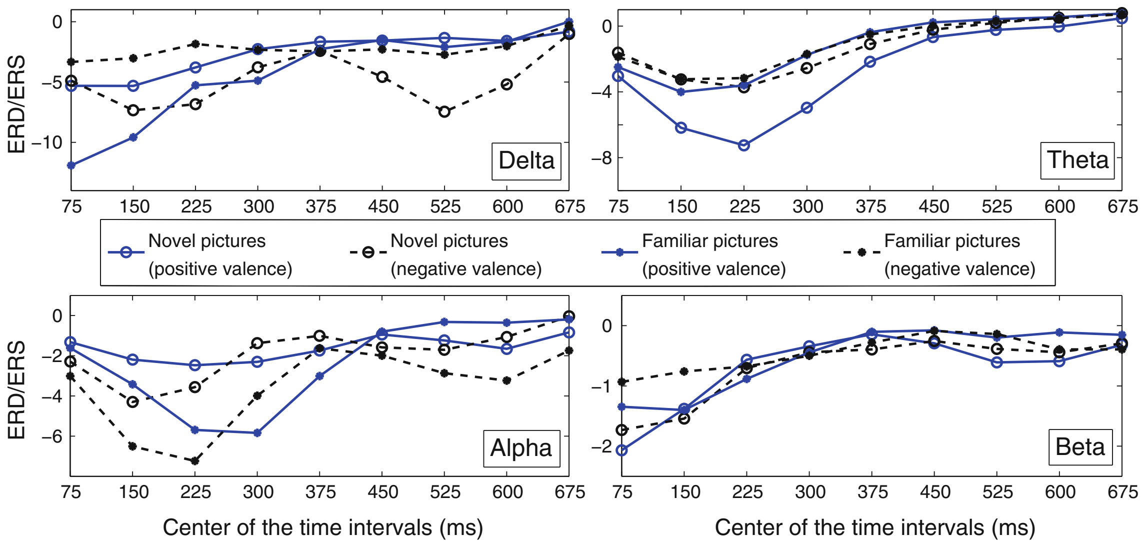

Fig. 2 ERD/ERS computed from frontal group in different frequency bands: delta (*up left*), theta (*up right*), alpha (*down left*), and beta (*down right*)

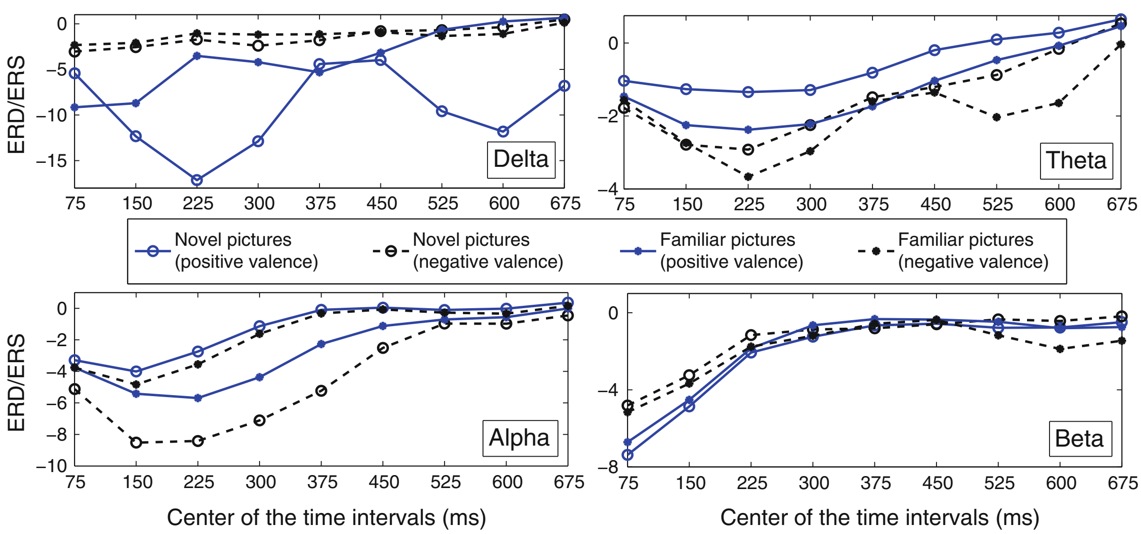

Fig. 3 ERD/ERS computed from parietooccipital group in different frequency bands: delta (*up left*), theta (*up right*), alpha (*down left*), and beta (*down right*)

improves. After k iterations, the vector $\mathbf{w}$ has $M \leftarrow M - k\tau$ elements, as well as the feature vector $\mathbf{x}$. The τ parameter can be selected according to the efficiency constraints or computational cost. In this study, since the number of initial features was not too high, a small step size (τ) was possible to employ. Experimentally, $\tau = 1$ and $\tau = 2$ were checked and no convergence of the algorithm could be achieved. Therefore, $\tau = 3$ was chosen as it was the minimum value that converged when all conditions were compared and showed an adequate resolution for discriminating the optimal number of features needed for reaching good accuracy values.

In this work, the feature vector is made up of 9 computed values e_i (ERD/ERS), measured for each of the 4 defined frequency bands in 2 scalp regions according to the grouping criteria, i.e., the initial number of features is $M = 9 \times 4 \times 2 = 72$. Furthermore, there will be two different classification tasks depending on which two scalp regions are considered: left and right (LR), or frontal and parietooccipital (FP). Each feature vector is labeled as $y = 1$ or $y = -1$ according to the sign of the valence of the stimulus, positive or negative, respectively. Specifically, the classification tasks are

- Task 1: Affective valence classification with novel pictures (NNC vs. NPC).
- Task 2: Affective valence classification with familiar pictures (FNC vs. FPC).

Note that our study does not aim at designing an optimal classification system, but at selecting the features that allow to distinguish in an accurate way the affective valence sign from only one ERP signal.

As the number of feature vectors belonging to each class is balanced, the accuracy yielded by the classifier can be considered a good measure of the system performance. At each iteration of the SVM-RFE feature elimination procedure, the classifier performance has been assessed using the leave-one-out cross-validation strategy, which consists of using all the samples in the dataset for training the system except one, which is used as test. This procedure is repeated S times, being S the number of samples in the classification task ($S = \text{subjects} \times \text{classes} = 26 \times 2 = 52$), after which a global value of accuracy is computed by counting the total number of correct decisions, out of S.

All the routines needed to apply the completed method were implemented in the software MATLAB (version R2011a, Mathworks, USA). In particular, tools from Bio-Informatics toolbox were required. The cost parameter of linear SVM was fixed on $C = 1$.

2.5 Statistical analysis (ANOVA)

An analysis of variance (repeated measures ANOVA) has been performed taking the whole set of features in order to find out whether statistically significant differences exist between both affective conditions (positive and negative valence) and both categories of pictures (novel and familiar). The different factors employed according to the condition are defined as follows: *habituation* (two levels: novel and familiar pictures); *valence* (two levels: negative and positive). Regarding the topography, the factor are defined as *position* (two levels: front/left and parietooccipital/right); *frequency band* (4 levels: δ, θ, α and β); and *time interval* (9 levels: from 0 to 750 ms, step = 100 ms, 50 % overlap). A significance level threshold of $p < 0.05$ has been considered, and Bonferroni's correction has been applied on post hoc tests. The Kolmogorov–Smirnov's test was previously performed for each sample, and from the results a normal distribution could be considered in every case. In addition, Levene's test was automatically applied for correcting the statistical significance if needed when homoscedasticity could not be assumed.

3 Results

Statistical methods have been widely used for extracting features and evaluating them accordingly to their significance in order to compare populations or different conditions of an experiment, being ANOVAs among the most popular ones. In order to compare this traditional method and find out to what extend the proposed SVM-RFE can complement or outperform it, both techniques have been implemented and discussed.

3.1 Preliminary statistical analysis

Only a few results have shown significant differences. In FP grouping, parietooccipital region and delta band are significantly different between conditions compared with the other regions and frequency bands, mainly using novel pictures. On the other hand, for LR grouping, right region in delta band is the most relevant area, according to the statistical significance ($p < 0.01$) with novel pictures, whereas left area in delta and alpha bands is the most important to determine the habituation phenomenon.

Table 1 shows the main effects obtained by applying an ANOVA to the database. Only the most interesting results are reported. Note that, by applying this statistical method, some conclusions could be extracted about the time course of ERD/ERS, reaching significant differences in its dynamics depending on the spectral range (see band × time effect). In general, when the time factor is exhaustively analyzed, only a clear significant difference ($p < 0.01$) in desynchronization values exists between

Table 1 Main effects obtained by applying a repeated measures ANOVA

Factors	FP grouping		LR grouping	
	F	p	F	p
hab	0.320	0.576	0.213	0.648
val	0.472	0.498	1.597	0.218
pos	0.304	0.586	0.362	0.553
band	2.859	**0.043**	3.857	**0.013**
time	2.239	**<0.001**	2.077	**<0.001**
hab × val	<0.001	0.986	0.088	0.769
hab × pos	0.276	0.604	0.396	0.535
val × pos	0.628	0.435	2.610	0.119
hab × band	0.437	0.727	1.560	0.206
val × band	1.617	0.193	1.104	0.353
pos × band	0.189	0.904	1.010	0.393
hab × time	0.646	0.739	1.711	0.098
val × time	1.580	0.133	1.790	0.081
pos × time	0.395	0.923	0.438	0.897
band × time	1.860	**0.008**	2.313	**<0.001**

The interactions between two factors (×) are also considered. F: Fisher's test; Bold values are statistically significant at $p < 0.05$. The factors are defined as hab: habituation, pos: position, band: frequency band, and time: time interval

FP frontal–parietooccipital grouping, *LR* left–right grouping early (up to 300 ms after stimulus onset) and late (from 375 ms after stimulus onset) time intervals (see also Figs. 2, 3).

3.2 SVM classifier performance and most relevant features

Figure 4 shows the accuracy rates obtained by applying a linear SVM classifier using separately the two regions defined in Sect. 2.4. In general, slightly higher values of accuracy are obtained by means of FP grouping in comparison with LR grouping. The accuracy rates using the ERPs linked to novel or familiar pictures visualization turns out to be the same. In all cases, the accuracy curve is not stable and shows local maximums and minimums, although the best results are obtained using around the 25 % of the total number of features (i.e., 16–20 features). The global maximum (Acc = 83 %) occurs when approximately the 25 % most relevant features are selected, with FP grouping and using for averaging only the ERPs from the familiar pictures trials (task 2).

It is convenient to separate both conditions (novel and familiar pictures) in different sets, since the classifier performance gets worse if they are mixed, i.e., forming two groups (positive and negative valences) joining in each one novel and familiar pictures and making only one classification task per condition and grouping method, for reaching the maximum Acc = 71 % and Acc = 65 % with FP and LR grouping, respectively. These values are lower than the highest accuracies obtained by using only one of the habituation-related conditions: Acc = 83 % (familiar condition) and Acc = 77 % (novel condition) with FP grouping; Acc = 74 % (familiar condition) and Acc = 75 % (novel condition) with LR grouping.

Table 2 compiles the most relevant time intervals and ERD/ERS per condition in each classification task and using the two grouping methods (FP and LR). The remaining features after running 15 iterations of the

Table 2 ERD/ERS most relevant time intervals in different frequency bands (δ: 0.5–4 Hz; θ: 4–7 Hz; α:7–12 Hz; β: 12–30 Hz) for discriminating affective valence

	FP grouping		LR grouping	
	Frontal	Parietooccipital	Left	Right
Task 1				
δ	[450, 525]	–	[150, 300, 450, 675]	[75, 300]
θ	[75, 150, 450]	[150, 300, 450, 675]	–	[450, 675]
α	[150, 300]	[75, 450]	[75, 300]	[150, 675]
β	[75, 300]	[225]	[600, 675]	[375, 675]
Task 2				
δ	[375]	[225, 300, 525, 675]	[150, 450, 675]	[225, 450]
θ	–	–	–	–
α	[600]	[150, 300, 375, 525]	[75, 300, 675]	[375, 600]
β	[300, 375]	[300, 375, 450, 675]	[300, 375, 450]	[150, 300, 675]

Up table: Task 1 (novel pictures condition); Bottom table: Task 2 (familiar pictures condition). Time intervals are named as the central instants (ms) in square brackets (*FP* frontal–parietooccipital grouping, *LR* left–right grouping)

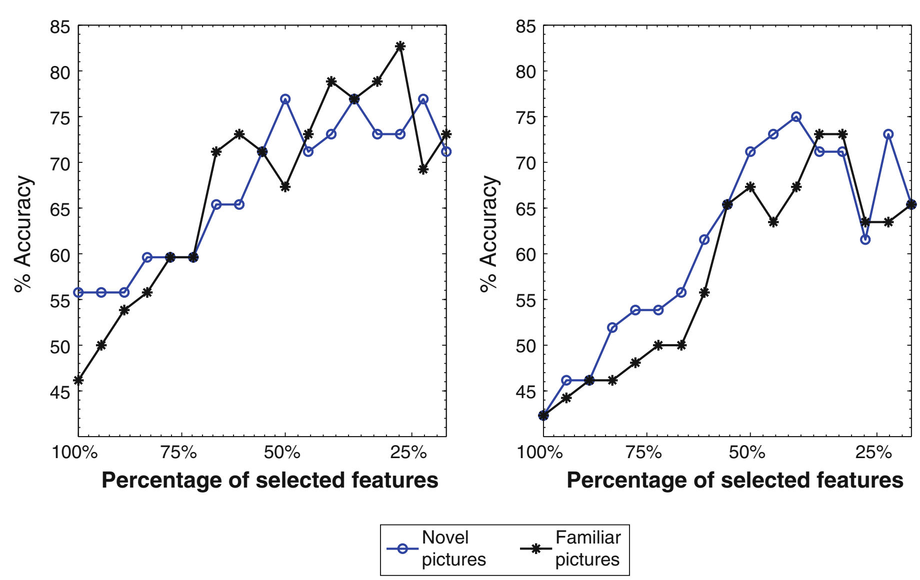

Fig. 4 Performance of linear SVM classifier *Left* FP grouping; *Right* LR grouping

algorithm SVM-RFE, where the maximal accuracy values were obtained (around the 22 % of the total number of features), have been chosen as the most relevant features. Parietooccipital area, closely linked to the P3 component in time domain, is especially relevant according to the large number of features selected (75 % of selected features are from parietooccipital region when familiar pictures are used).

4 Discussion

This work proposes a methodology to study the ERPs for finding out how the affective valence processing after visualizing high arousal pictures is affected by the habituation phenomenon, specifically in young adult women. The ERPs are estimated with less number of trials than it is usual and using different criteria for grouping the EEG channels, resulting then only one ERP for one specific brain region. Afterward, by computing the ERD/ERS in different moments after stimulus presentation and in different ranges of frequencies, the most relevant features are found for discriminating between positive and negative valence conditions. For this purpose, a wrapper supervised classification method, named SVM-RFE, was applied on the set of extracted features, yielding 83 % accuracy in distinguishing the two conditions (positive and negative) within familiar and novel images. This study had never been carried out before and is able to provide additional information to ANOVAs.

Statistical methods like ANOVA are useful in order to contrast hypotheses about the similarity of mean values in measures taken from different experimental conditions or groups. Nevertheless, they do not always provide statistically significant results when they involve corrections, in the case of multiple comparisons, that might obscure and make more difficult the extraction of clear conclusions. In this work, even though significant differences in time intervals have been found (which coincide with other previously reported works [4]), relevance of the novelty factor in stimuli cannot be properly drawn. For some purposes such as BCI or online applications, only one signal is provided and the significant features must be chosen without instantaneous comparisons with other complementary signals. For this reason, the use of SVM-RFE, which considers the initial set of features as a whole without comparing directly with any other antithetical observation, is a good choice to complement the information provided by ANOVA.

While the channel grouping method influences in the performance of the classifier, the habituation process neither increases nor decreases the accuracy values in any classification task. Generally, results reported previously in the literature on affective valence detection hardly exceed Acc = 70 % on average [16, 27], being worse than arousal classification, in terms of accuracy rates.

The equivalent contribution of each hemisphere is patent in this study (see LR grouping results in Table 2, where the number of selected features is very similar in both hemispheres). Asymmetry, which is observed in other studies about emotions [11], is not clearly found between hemispheres when frontal and parietooccipital areas are joined in the same (left or right) cluster. Furthermore, the spread of the relevant ERDs/ERSs, across all frequency bands, does not clearly support an asymmetrical contribution of alpha waves from both hemispheres, which has been reported in other research works [11].

In general, comparing the graphics of Fig. 2 which represents the grand averaged extracted features for the frontal group, the visual intuition matches up with the results of the algorithm. For instance, the time interval "525" in delta band is relevant according to SVM-RFE, and it is really different in the graphics if NNC and NPC are compared. However, it is neither possible nor advisable to be guided only by average illustrations for extracting conclusions about the significance or relevance of time intervals in ERP, since occasional artifacts, for instance the activity from muscles in bordering electrodes or border effects due to the windowing, could alter the grand averages. The time interval "225" seems to be significantly different in theta band between NNC and NPC, but it is not selected as one of the most relevant by SVM-RFE. In regard to Fig. 3, the greatest differences between conditions seem to appear in delta band using novel pictures, nevertheless not a single feature is selected from this band.

Regarding the most relevant frequency bands for discriminating the affective valence in EEG signals, they are clearly dependent on the habituation phenomenon. For example, ERD/ERS in delta band is not especially relevant in valence classification with novel pictures if FP grouping is used, maybe due to some contribution from ocular artifacts in frontal areas, but it is relevant in LR grouping (see Table 2). Results in theta band are remarkable: this band has not any contribution to the most relevant features for valence classification using familiar pictures, but it has a high contribution if novel pictures and FP grouping are used. In the previous research works, it has been demonstrated that theta band is linked to affective valence processing [1], allowing to distinguish between emotional and neutral states, either with explicit and implicit emotions. The conclusions extracted by the study of Mu et al. [17] are especially interesting to link to the habituation phenomena observed in our results, since they relate theta and alpha oscillations to empathy for pain. Note that the negative pictures chosen from the IAPS mainly show injuries or mutilated people. This way, the higher relevance of theta band for distinguishing the affective valence using novel pictures can be influenced by the empathy for pain, whereas the repetition of the same pictures could favor the reduction of that empathetic response. Nonetheless, this idea must be taken carefully, since it might contradict a generally accepted conclusion about the stability of ERP elicited by negative stimuli.

To summarize, the obtained results conclude that, for classifying the affective valence induced by picture visualization, parietooccipital region is the most discriminant scalp area and theta band synchronization is affected by habituation phenomenon. Nevertheless, it is bold to relate, in a restricted sense, only one frequency band to only one specific phenomenon, even more when a subjective affective valence is dealt. Although negative stimuli are usually less prone to habituation process than positive stimuli, it is convenient to use the two options (novel or familiar pictures) separately, in order to better classify between negative and positive valence, in BCI or artificial intelligence applications.

Acknowledgments This work is partially funded by FEDER through the Operational Program Competitiveness Factors—COMPETE and by National Funds through FCT—Foundation for Science and Technology in the context of the project FCOMP-01-0124-FEDER-022682 (FCT reference PEst-C/EEI/UI0127/2011).

References

1. Aftanas LI, Varlamow AA, Pavlov SV, Makhnev VP, Reva NV (2001) Affective picture processing: event-related synchronization within individually defined human theta band is modulated by valence dimension. Neurosci Lett 303:115–118
2. Bradley MM, Lang PJ, Cuthbert BN (1993) Emotion, novelty, and the startle reflex: habituation in humans. Behav Neurosci 107:970–980
3. Carretié L, Hinojosa JA, Mercado F (2003) Cerebral patterns of attentional habituation to emotional visual stimuli. Psychophysiology 40:381–388
4. Cesarei AD, Codispoti M (2011) Affective modulation of the LPP and alpha-ERD during picture viewing. Psychophysiology 48:1397–1404
5. Codispoti M, Ferrari V, Bradley MM (2007) Repetition and event-related potentials: distinguishing early and late processes in affective picture perception. J Cogn Neurosci 19:577–586
6. Ferrari V, Bradley MM, Codispoti M, Lang PJ (2011) Repetitive exposure: brain and reflex measures of emotion and attention. Psychophysiology 48:515–522
7. Frantzidis CA, Bratsas C, Papadelis CL, Konstantinidis E, Pappas C, Bamidis PD (2010) Toward emotion aware computing: an integrated approach using multichannel neurophysiological recordings and affective visual stimuli. IEEE Trans Inf Technol B 14:589–597
8. Fuentes A, Farina D, Dremstrup K (2010) Comparison of feature selection and classification methods for a brain–computer interface driven by non-motor imagery. Med Biol Eng Comput 48:123–132
9. Guyon I, Weston J, Barnhill S, Vapnik V (2002) Gene selection for cancer classification using support vector machines. Mach Learn 46:389–422
10. Hidalgo-Muñoz AR, López MM, Santos IM, Pereira AT, Vázquez-Marrufo M, Galvao-Carmona A, Tomé AM (2013) Application of SVM-RFE on EEG signals for detecting the most relevant scalp regions linked to affective valence processing. Expert Syst Appl 40:2102–2108
11. Huang D, Guan C, Ang KK, Zhang H, Pan Y (2012) Asymmetric spatial pattern for EEG-based emotion detection. In: World Congress on Computational Intelligence. WCCI 2012. IEEE
12. Klados MA, Frantzidis C, Vivas AB, Papadelis C, Lithari C, Pappas C, Bamidis PD (2009) A framework combining delta event-related oscillations (EROs) and synchronisation effects (ERD/ERS) to study emotional processing. Comput Intell Neurosci 2009, Article ID 549419. doi:10.1155/2009/549419
13. Lang PJ, Bradley MM, Cuthbert BN (2008) International affective picture system (IAPS): instruction manual and affective ratings. Technical Report A-8, The Center for Research in Psychophysiology, University of Florida
14. Lithari C, Frantzidis C, Papadelis C, Vivas AB, Klados M, Kourtidou-Papadeli C, Pappas C, Ioannides A, Bamidis P (2010) Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. Brain Topogr 23:27–40
15. López MM, Ramírez J, Górriz JM, Álvarez I, Salas-Gonzalez D, Chaves R (2009) SVM-based CAD system for early detection of the Alzheimer's disease using kernel PCA and LDA. Neurosci Lett 464:233–238
16. Macas M, Vavrecka M, Gerla V, Lhotska, L (2009) Classification of the emotional states based on the EEG signal processing. In: Proceedings of the 9th International Conference on Information Technology and Applications in Biomedicine, Larnaca, Cyprus, pp 1–4
17. Mu Y, Fan Y, Mao L, Han S (2008) Event-related theta and alpha oscillations mediate empathy for pain. Brain Res 1234: 128–136
18. Nielen M, Heslenfeld D, Heinen K, Strien JV, Witter M, Jonker C, Veltman D (2009) Distinct brain systems underlie the processing of valence and arousal of affective pictures. Brain Cogn 71:387–396
19. Olofsson JK, Nordin S, Sequeira H, Polich J (2008) Affective picture processing: an integrative review of ERP findings. Biol Psychol 77:247–265
20. Pfurtscheller G, Da Silva FHL (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 110:1842–1857
21. Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton D, Colombo J, Coppola G, Geyer MA, Glanzman DL, Marsland S, McSweeney F, Wilson DA, Wu CF, Thompson RF (2009) Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiol Learn Mem 92:135–138
22. Rozenkrants B, Olofsson JK, Polich J (2008) Affective visual event- related potentials: arousal, valence, and repetition effects for normal and distorted pictures. Int J Psychophysiol 67:114–123
23. Shawe-Taylor J, Sun S (2011) A review of optimization methodologies in support vector machines. Neurocomputing 74(17):3609–3618
24. Sun S (2008) The extreme energy ratio criterion for EEG feature extraction. Lect Notes Comput Sci 5164:919–928
25. Sun S (2010) Extreme energy difference for feature extraction of EEG signals. Expert Syst Appl 37(6):4350–4357
26. Sun S, Zhang C (2006) Adaptive feature extraction for EEG signal classification. Med Biol Eng Comput 44:931–935
27. Tomé AM, Hidalgo-Muñoz AR, López MM, Teixeira AR, Santos IM, Pereira AT, Vázquez-Marrufo M, Lang EW (2013) Feature

extraction and classification of biosignals: emotion valence detection from EEG signals. In: proceedings of International Conference on Bio-inspired Systems and Signal. BIOSIGNALS 2013, pp 54–60
28. Weierich MR, Wright CI, Negreira A, Dickerson BC, Barrett LF (2010) Novelty as a dimension in the affective brain. Neuroimage 49:2871–2878
29. Xu H, Plataniotis KN (2012) Affect recognition using EEG signal. In: Proceedings of the 14th International Workshop on Multimedia Signal Processing. Banff, Canada, pp 299–304
30. Yoon HJ, Chung SY (2011) EEG spectral analysis in valence and arousal dimensions of emotion. In: Proceedings of the 11th International Conference on Control, Automation and Systems, Kintex, Gyeonggi-do, Korea, pp 1319–1322