Feature Extraction and Classification of Biosignals
Emotion Valence Detection from EEG Signals

A. M. Tomé[1], A. R. Hidalgo-Muñoz[2], M. M. López[3], A. R. Teixeira[3], I. M. Santos[4], A. T. Pereira[4], M. Vázquez-Marrufo[2] and E. W. Lang[5]

[1]*DETI/IEETA, University of Aveiro, 3810-193 Aveiro, Portugal*

[2]*Dept. Experimental Psychology, University of Seville, 41018 Seville, Spain*

[3]*IEETA, University of Aveiro, 3810-193 Aveiro, Portugal*

[4]*Dept Ciências de Educação, University of Aveiro, 3810-193 Aveiro, Portugal*

[5]*Biophysics, CIML Group, University of Regensburg, 93040 Regensburg, Germany*

Keywords: Valence Detection, Random Forest, ERD/ERS.

Abstract: In this work a valence recognition system based on electroencephalograms is presented. The performance of the system is evaluated for two settings: single subjects (intra-subject) and between subjects (inter-subject). The feature extraction is based on measures of relative energies computed in short time intervals and certain frequency bands. The feature extraction is performed either on signals averaged over an ensemble of trials or on single-trial response signals. The subsequent classification stage is based on an ensemble classifier, i. e. a random forest of tree classifiers. The classification is performed considering the ensemble average responses of all subjects (inter-subject) or considering the single-trial responses of single subjects (intra-subject). Applying a proper importance measure of the classifier, feature elimination has been used to identify the most relevant features of the decision making.

1 INTRODUCTION

During the last decades, 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 recognition, voice recognition and electrophysiology - based systems (Calvo and D'Mello, 2010). Concerning the origin of the signals of the latter systems, they can be divided into two categories: those originating from the peripheral nervous system (e.g. heart rate, Electromyogram - EMG, galvanic skin resistance-GSR), and those coming from the central nervous system (e.g. Electroencephalograms-EEG). Traditionally, EEG-based technology has been used in medical applications but nowadays it is spreading to other areas like entertainment and brain-computer interfaces (BCI). With the emergence of wearable and portable devices, developing systems based on EEG signals attracted much interest. Therefore, with the availability of vast amounts of digital data, there is an increasing interest in the development of machine learning software applications. Emotion recognition systems, dealing with biological signals, exhibit performances ranging from 40% to 90% depending on the number of emotion categories of the study. However, it is not easy to compare them since they differ in the way emotions are elicited, and in the underlying model of emotions (e.g. emotional categories). Psychologists represent emotions in a 2D valence/arousal space (Bradley and Lang, 2007). By dividing the valence (horizontal axis) - arousal (vertical axis) space into four quadrants, several discrete emotions are usually identified (Russell, 1980). The most widely used categories are the following : *Joy* (high valence , high arousal); *Pleasure* (high valence, low arousal); *Anger* (Low valence, High arousal); *Sadness* (low valence, low arousal). Some studies include a fifth category assigned as *Neutral* which is represented in the region close to the origin of the 2D valence/arousal space. Some studies concentrated on one of the dimensions of the space like identifying the arousal intensity (high versus low) or the valence (low/ negative versus high /positive), and eventually a third class neutral state. Normally, emotions are elicited by (i) presenting an

M. Tomé A., R. Hidalgo-Muñoz A., M. López M., R. Teixeira A., M. Santos I., T. Pereira A., Vázquez-Marrufo M. and W. Lang E..
Feature Extraction and Classification of Biosignals - Emotion Valence Detection from EEG Signals.
DOI: 10.5220/0004233100540060
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* (BIOSIGNALS-2013), pages 54-60
ISBN: 978-989-8565-36-5

external stimulus (picture, sound, word or videos) related to different emotions at some predefined interval, or by (ii) simply asking subjects to imagine different kinds of emotions. Concerning external visual stimuli, one may resort to the International Affective Picture System (IAPS) collection which is freely available (Lang et al., 2008), and is composed of pictures classified by a large number of participants in terms of *Arousal* and *Valence*. The picture set is widely used in experimental psychology as well as in automatic emotion recognition systems. Most of the work in automatic recognition can be considered pilot studies where all experiments are conducted under laboratory settings where experiments are prepared to induce emotions. But like in any other classification system, it is needed to establish which signals and how they are acquired, extract relevant features from these input signals, and finally train a classifier. In this work, a valence recognition system is presented, the feature extraction module of which is inspired on brain studies, and its classification module is based on a random forest of decision trees. The latter module is also applied recursively to achieve feature elimination. Moreover, the classification performance is measured with respect to inter- and intra-subject classification. To achieve such goals, different types of signals are applied as inputs to the feature extraction module: either single-trial or ensemble average signals.

1.1 Classification Systems and Emotion

The pioneering work of Picard (Picard et al., 2001) 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 (Picard et al., 2001). In (Haag et al., 2004), using the IAPS data set as stimulus reportoir, 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 consider that the estimation of the valence value (63.8%) is a much harder task than the estimation of arousal (89.3%). In (Kim et al., 2004), a study with 50 participants, aged from seven to eight years old, is presented. The visual stimulation with the IAPS data set was considered insufficient, hence they proposed a sophisticated scenario to elicit emotions. It consisted of visual stimuli using controlled illumination, and additional auditory stimuli using background music. Simultaneously, an actress narrated a story (with emotional contents, like sadness) that was carefully prepared to evoke the sympathy of the subjects. The latter were also requested to look at a toy in front of them, and it seemed as if the toy was telling the sad story to the subjects. Only peripheral biological signals were recorded, and the measured features were the input of a classification scheme based on a support vector machine (SVM). The results showed accuracies of 78.4% and 61% for 3 and 4 categories of different emotions, respectively. In (Schaaff and Schultz, 2009), the data collection was performed with stimulus pictures taken from the IAPS repository thus inducing three emotional states 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 et al., 2009), where the EEG data was collected from 23 subjects during an affective picture stimuli 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 (Frantzidis et al., 2010), four emotional categories of the arousal/valence space of the IAPS picture set were used to elicit emotions of 28 participants and their EEG signals were recorded. The ensemble average was 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.

1.2 Event Related Potentials and Emotion

Most of the recognition systems referred above extract features in segments of the signal defined after the stimulus presentation. Those features were found relevant in brain studies and are generally addressed in studies of event-related potentials (ERP). ERPs represent transient components in the electroencephalogram (EEG) generated in response to a stimulus, e.g. a visual or auditory stimulus. Studies of event-related potentials deal with signals that represent different levels of analysis: signals from single-trials, ensemble averaged signals where the ensemble encompasses several trials, and signals resulting from a grand-average over different trials as well as subjects. The segment of the time series containing the single-trial response signal is centered on the stimulus: t_i (negative value) before and t_f (positive value) after stimulus. The ensemble average, over trials of one subject, eliminates the spontaneous activity of brain maintaining the activity that is phase-locked with stimulus. And the grand-average is the average, over participants, of ensemble averages and it is used mostly to illustrate the outcomes of the study. In experimental psychology studies, ERP is usually the ensemble average computed with all single-trials belonging to one condition (stimulus type). Those works show that the event-related potentials (ERP) have characteristics (amplitude and latency) of the early waves which change according to the nature of the stimuli (Olofsson et al., 2008). Other investigations studied the effect of the stimulus in the characteristics frequency bands. Hence, these measures reflect changes in α -, β -, θ - or δ - bands. One of the most popular, simple and reliable measures is the event related desychronization/synchronization (ERD/ERS). It represents a relative decrease or increase in the power content in time - intervals defined after the stimulus onset when compared to a reference interval defined before the stimulus onset (Klados et al., 2009). Usually this measure is computed for the different characteristic bands of the EEG (Klados et al., 2009).

2 METHODOLOGY

In this work a valence recognition system is presented. The performance of the system is evaluated for both single subjects (intra-subject) and between subjects (inter-subject). The feature extraction is based on ERD/ERS measures computed in short intervals. The subsequent classification stage is based on an ensemble classifier, i. e. a random forest of tree classifiers. The feature extraction is performed either on signals averaged over an ensemble of trials or on single-trial response signals. Accordingly, the classification is performed considering the ensemble average responses of all subjects (inter-subject) or considering the single-trial responses of single subjects (intra-subject). Furthermore, feature elimination, applying a proper importance measure of the classifier, has been used to identify the most relevant features of the decision making.

2.1 The Dataset

A total of 26 female volunteers participated in the study. A total of 21 channels of EEG, positioned according to the $10-20$ system, and 2 Electrooculograms (EOG) channels (vertical and horizontal) were sampled at $1kHz$ and stored. The signals were recorded while the volunteers were viewing pictures selected from the IAPS picture repository. A total of 24 high-arousal images, corresponding to an arousal score $s > 6$, with positive valences ($v = 7.29 \pm 0.65$) and negative valences ($v = 1.47 \pm 0.24$) were selected. Each image was presented three times in a pseudo-random order and each trial lasted $3500ms$: during the first $750ms$, a fixation cross was presented, then one of the images was shown during $500ms$, and finally a black screen followed for a period of $2250ms$. The signals were pre-processed (filtered, eye-movement corrected, baseline compensated and segmented into epochs) using the NeuroScan software package. The single-trial signal length is $950ms$, with $150ms$ before the stimulus onset.

2.2 Feature Extraction

The features are extracted from the segmented signals (either ensemble averaged or single-trial), measuring the desychronization/synchronization (ERD/ERS) in four frequency bands. Then the signals are filtered by four 4th order bandpass Butterworth filters. The four characteristic pass-bands are defined as: $\delta - band$: $0.5-4Hz$, $\theta - band : 4-7Hz$, $\alpha - band : 8-12Hz$ and $\beta - band : 13-30Hz$. The $K = 4$ filters are applied following a zero-phase forward and reverse digital filter methodology not including any transient (see *filtfilt* MATLAB function (Mathworks, 2012)). For each filtered signal, the ERD/ERS is estimated in $I = 9$ intervals following the stimulus onset and with a duration of $150ms$ and 50% of overlap between consecutive intervals. The reference interval corresponds to the $150ms$ pre-stimulus period. For each interval, the ERD/ERS is defined as

$$f_{ik} = \frac{E_{rk}-E_{ik}}{E_{rk}} = 1 - \frac{E_{ik}}{E_{rk}} \quad i = 1,2\ldots9; k = 1,\ldots4$$

where E_{rk} represents the energy within the reference interval, while E_{ik} is the energy in the $i-th$ interval after stimulus in the $k-th$ band. Note that when $E_{rk} > E_{ik}$, the 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. In this work, energies E_{ik} are computed by adding up instantaneous energies within each of the $I = 9$ intervals of $150ms$ duration. The energy E_{rk} is estimated in an interval of $150ms$ duration defined in the pre-stimulus period. Figure 1 represents the features computed for two ensemble signals of channel $F7$ of one subject. In summary, each valence condition can be characterized by f_{ikc}, where i stands for the time interval, k for the characteristic

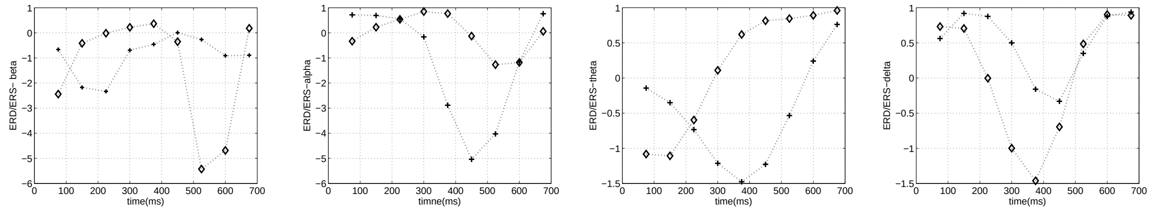

Figure 1: Features in ensemble average signals of $F7$ channel of one subject: negative valence ($\diamond$) versus positive valence ($+$).

band and c for the channel. A total of $M = I \times K \times C = 9 \times 4 \times 21 = 756$ features is computed for the multichannel segments related with one condition. Following, the features f_{ikc} will be concatenated into a feature vector with components f_m, $m = 1, \ldots, M = 756$.

2.3 Random Forest

Random forest is an ensemble classifier where training is based on bootstrapping techniques. The random forest algorithm, developed by Breiman (Breiman, 2001), is a set of binary decision trees, each performing a classification and the final decision is 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 $\omega_c, c = 1 \ldots C$ are the labels of the classes, the Gini index of node i is given by

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

where the $P(\omega_c)$ 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 and the change in impurity is computed as

$$\triangle G(i) = G(i) - \Big(\frac{n_l}{n} G(i_l) + \frac{n_r}{n} G(i_r)\Big)$$

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. The grown trees are not pruned.

The main steps of the algorithm are

1. Given a data set $\mathcal{T}$ 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 $\mathcal{T}_t$ with N examples by sampling with replacement the original data set. The out-of-bag data set O_t is formed with the remaining examples of $\mathcal{T}$ not belonging to $\mathcal{T}_t$.
   (b) Perform the split of node i by testing one of the $L = \lfloor \sqrt{F} \rfloor$ randomly selected features.
   (c) Repeat step 2b up to 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 $t \neq T$.

After training, the importance r_m of each feature f_m in the ensemble of trees can be computed by adding the values of $\triangle G(i)$ of all nodes i where the feature f_m is used to perform a split. Sorting the values r_m by decreasing order, it is possible to identify the relative importance of the features. In this work $T = 500$ decision trees were employed.

2.4 Classification and Feature Elimination

In (Guyon et al., 2002) a recursive feature elimination scheme is proposed based on the values of the parameters of the classifier. In this work a similar strategy using the variable importance r_m is applied according to the following scheme:

1. Initialize: create a set of indices $\mathcal{M} = \{1, 2, \ldots M\}$ relative to the available features and set $F \equiv M$
2. Organize data set $\mathcal{X}$ by forming the feature vectors with the feature values whose index is in set $\mathcal{M}$
3. Compute the accuracy of the classifier using either leave-one-out or k-fold cross-validation.
4. Compute the global model of the classifier using the complete data set $\mathcal{X}$.
5. Compute r_m of the features set and eliminate from set $\mathcal{M}$ 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 to 6 while set $\mathcal{M}$ is not empty.

The leave-one-out strategy was followed in the intra-subject experiments and in inter-subject experiments when the ensemble averages were computed with all available trials of each subject and each condition. The k-fold cross-validation strategy was used in inter-subject experiments otherwise. Each fold is formed with the data of each subject, e.g. the classifier is trained with features extracted from 25 subjects leaving the data of the remaining subject to estimate the accuracy.

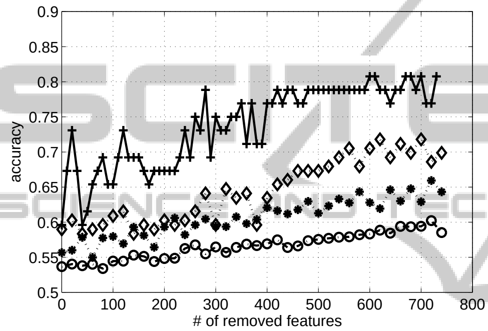

Figure 2: The inter-subject accuracy versus feature selection. Features extracted from Ensemble-average signals with: $+$ with at least 30 single-trials, $\diamond$ 10 consecutive single-trials;$*$ of 3 consecutive single-trials ;and O single-trial signal.

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. But this training and test procedure is repeated such that all the elements of the data set are used as test set. Therefore the accuracy is the proportion of correct decisions taken by classifier during the execution of the leave-one-out loop strategy. In the intra-subject study the accuracy can be estimated as the average of the estimates within each subject.

3 RESULTS AND DISCUSSION

The system was implemented in MATLAB using also some facilities of open source software tools like EEGLAB (Delorme and Makeig, 2004) and the random forest package (Jaiantilal, 2010). Considering feature elimination and the concomitant number of relevant features, as can be seen from fig. 2 and fig. 3, the accuracy of the classifier improves with a decreasing number of relevant features in both an inter-subject or an intra-subject classification strategy. In every case, the accuracy achieves 80% when the classifiers have less than 100 relevant features as input.

3.1 Inter-subject Classification

Figure 2 shows the accuracy versus the number of features eliminated. The accuracy was computed with a leave-one-out strategy and a total of 52 feature vectors were involved. The highest accuracy is achieved having as input the ensemble average of all trials. An average accuracy of 79% is achieved if roughly 500 irrelevant features are removed from the input feature set. The other traces represent the mean accuracy when the ensemble averages are computed with $3,5,10$ consecutive trials for each condition meaning that the number of feature vectors for each subject is roughly $24,14,6$, respectively. Note that the differences in accuracy between the various cases considered might not be statistically significant rather appear as a consequence of the sample size of the test sets. However notice that the curves follow a similar trend corroborating the positive effect on the decision making of eliminating irrelevant features, i. e., distracting information.

The tables 1 and 2 describe the spatial and temporal location of the relevant selected features when the input of the classifier is the data set formed by 52 feature vectors. These feature vectors represent the ensemble average positive and negative response of all volunteers investigated. Concerning spacial locations, the largest number of features happens to occur in the frontal and parietal regions of the brain. Considering the localization of the response in time, most of the features display *Medium* and *Long* latencies. These results confirm related brain studies performed with ensemble average signals (Olofsson et al., 2008).

Table 1: Space Localization of the 36 selected features within each band: frontal (Fr), central-temporal (CT) and parietal-occipital (PO).

Channels	Beta	Alpha	Theta	Delta	Total
Fr.	7	2	4	5	18
CT	6	0	3	0	9
PO	5	2	0	2	9

3.2 Intra-subject Classification

Figure 3 shows the mean accuracy when the classifier is trained with data of one subject. The features were extracted from single-trial signals as described before. The training set for each subject is formed by a total

Table 2: Time Localization of the 36 selected features within each band. Time intervals: Short ($i = \{1,2\}$), Medium ($i = \{3,4\}$), Long I ($i = \{5,6\}$) and Long II ($i = \{7,8,9\}$).

Time	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

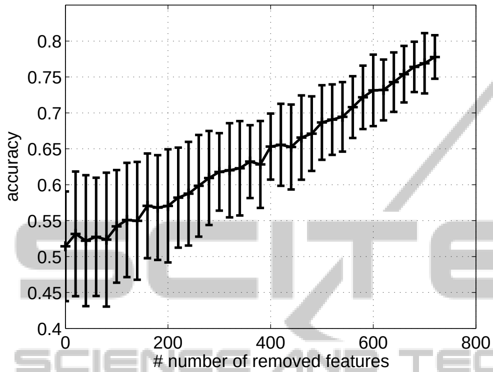

Figure 3: Average and standard deviation accuracy of intra-subjects accuracies versus number of features removed. The last point corresponds to a removal of 720 features.

of $65-72$ single trials for both classes of emotions. Again a leave-one-out strategy was employed.

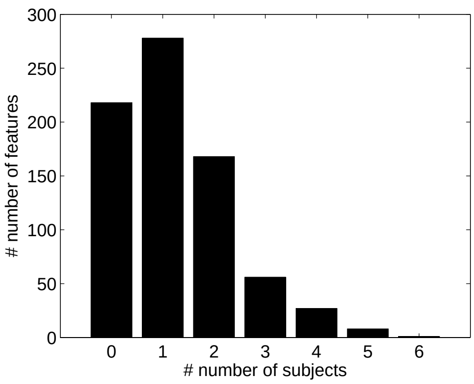

Figure 4: Within the 36 features selected from each individual training the histogram counts the number of times a feature was selected.

A comparison of the outcomes of the individual training sessions with respect to the features that remain after performing the same number of iterations reveals a large inter-subject variability. Figure 4 displays this comparison when all individual training sessions had 36 features as input. It can be seen that 218 features turn out to be completely irrelevant and have been eliminated from all classification sessions. Another 275 features appear as relevant features for the decision making in one of the subjects under study. Remarkably, only one feature appears consistently as a relevant feature in at least 6 out of 26 subjects confirming a high inter-subject variability. A similar conclusion has been drawn in (Macas et al., 2009) by using a feature selection block before performing classification. However notice that a comparable accuracy is achieved whether decision making is based on a set of 52 feature vectors which represent ensemble averages over trials and subjects or whether decision making is based on training classifiers individually with $65-72$ feature vectors for each subject. As can be seen in table 3, on average, the most relevant features are appearing again in the frontal region. Note, however, that this time the parietal-occipital region seems more relevant than the central-temporal region.

Table 3: Spatial location versus Frequency: frontal (Fr), central-temporal (CT) and parietal-occipital (PO). Average and standard deviation values of number of features within the 36 selected on each subject.

Ch.	Beta	Alpha	Theta	Delta
Fr.	4.6 ± 2.9	3.5 ± 4.1	2.9 ± 2.6	4.0 ± 3.1
CT	3.0 ± 2.9	2.5 ± 2.2	1.2 ± 1.6	1.6 ± 2.3
PO	2.8 ± 2.3	2.8 ± 2.7	2.9 ± 2.7	4.1 ± 3.4

4 CONCLUSIONS

A valence recognition system has been presented and applied to EEG signals. The latter were recorded from volunteers subjected to emotional states elicited by visual stimuli drawn from IAPS repository. The recognition system encompasses a feature extraction stage and a classification module including feature elimination. A cohort of 26 female volunteers (age $18-62$ years; mean=24.19; std=10.46) has been investigated. Feature extraction was based on an inter-subject and an intra-subject methodology. Both methodologies showed similar performance with regard to the accuracy of the random forest classifier. However from the related Gini index measuring feature importance no consistent set of features could be identified supporting the decision making. This points towards a large biological variability of the set of relevant features corresponding to the valence of the emotional states involved. The classification accuracy achieved compares well with or is even superior to related systems reported in literature.

Although inter-subject and intra-subject methodologies show a similar performance they yet have different application scenarios. The inter-subject is mostly suitable for off-line applications like brain studies in order to complement the statistical methods. For instance, in (Hidalgo-Munoz et al., 2012) an SVM-RFE scheme was applied to identify scalp spectral dynamics linked with the affective valence processing. While intra-subject might be interesting for personalized studies, where subjects need to be followed over a couple of sessions. Because of the biologically variability observed intra-subject studies cannot generalized easily across a cohort of subjects.

ACKNOWLEDGEMENTS

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). The Financial support by the DAAD - FCT is also gratefully acknowledged.

REFERENCES

Bradley, M. M. and Lang, P. J. (2007). The International Affective Picture System (IAPS) in the Study of Emotion and Attention. In Bradley, M. M. and Lang, P. J., editors, *Handbook of emotion elicitation and assessment*, Series in Affective Science, chapter 2, pages 29–46. Oxford University Press.

Breiman, L. (2001). Random forests. *Mach. Learn.*, 45:5–32.

Calvo, R. A. and D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing*, 1(1):18–37.

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

Frantzidis, C. A., Bratsas, C., Klados, M. A., Konstantinidis, E., Lithari, C. D., Vivas, A. B., Papadelis, C. L., Kaldoudi, E., Pappas, C., and Bamidis, P. D. (2010). On the classification of emotional biosignals evoked while viewing affective pictures: An integrated data-mining-based approach for healthcare applications. *Information Technology in Biomedicine, IEEE Transactions on*, 14(2):309–318.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.*, 46(1-3):389–422.

Haag, A., Goronzy, S., Schaich, P., and Williams, J. (2004). Emotion recognition using bio-sensors: First steps towards an automatic system. In André, E., Dybkjæ r, L., Minker, W., and Heisterkamp, P., editors, *Affective Dialogue Systems*, volume i, pages 36–48. Springer.

Hidalgo-Munoz, A., López, M., Santos, I., Pereira, A., Vázquez-Marrufo, M., Galvao-Carmona, A., and Tomé, A.M. (2012). Application of SVM-RFE on EEG signals for detecting the most relevant scalp regions linked to affective valence processing. *Expert Systems with Applications*, in press(0).

Jaiantilal, A. (2010). http://code.google.com/p/randomforest-matlab/

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

Klados, M. A., Frantzidis, C., Vivas, A. B., Papadelis, C., Lithari, C., Pappas, C., and Bamidis, P. D. (2009). A Framework Combining Delta Event-Related Oscillations (EROs) and Synchronisation Effects (ERD/ERS) to Study Emotional Processing. *Computational intelligence and neuroscience*, 2009:549419.

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

Macas, M., Vavrecka, M., Gerla, V., and Lhotska, L. (2009). Classification of the emotional states based on the EEG signal processing. In *2009 9th International Conference on Information Technology and Applications in Biomedicine*, number November, pages 1–4. IEEE.

Mathworks (2012). http://www.mathworks.com/help/signal /ref/filtfilt.html

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

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

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

Schaaff, K. and Schultz, T. (2009). Towards emotion recognition from electroencephalographic signals. In *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pages 1–6. IEEE.