# Classification of the emotional states based on the EEG signal processing

Martin Macaš, Michal Vavrečka, Václav Gerla and Lenka Lhotská, *Member, IEEE*

**Abstract — The paper proposes a method for the classification of EEG signal based on machine learning methods. We analyzed the data from an EEG experiment consisting of affective picture stimuli presentation, and tested automatic recognition of the individual emotional states from the EEG signal using Bayes classifier. The mean accuracy was about 75 percent, but we were not able to select universal features for classification of all subjects, because of inter-individual differences in the signal. We also identified correlation between the classification error and the extroversion-introversion personality trait measured by EPQ-R test. Introverts have lower excitation threshold so we are able to detect the differences in their EEG activity with better accuracy. Furthermore, the use of Kohonen's self-organizing map for visualization is suggested and demonstrated on one subject.**

Index terms: feature selection, SOM, Bayes classifier, emotional states, EEG

## I. Introduction

RECENT development in the area of EEG signal processing allows us to classify not only the basic conscious states as sleeping stages or the level of vigilance but also to focus on the more specific aspects of the human cognitive processing. One of the subfield in this area deals with identification of specific affects and emotions and their correlates in the biologic signals.

There are several methods to measure brain activity (f MRI, EEG, PET) in the area of cognitive neuroscience, but we will focus on the EEG signal and its frequency analysis. The well known research line of R. J. Davidson [1]-[3] uncovers regularities between the affective stimulation and the subject's inter hemispheric asymmetric responses in the alpha band wave, especially in the frontal and prefrontal brain areas. These findings were confirmed in the studies of Coan and Allen [4], [5]. They found the correlates of the emotional stimuli in the frontal asymmetries and consolidated some methodological problems with the experimental design. They also overview the history of the emotional processing research and summarize main differences among the researches [5].

Although the results of the mentioned studies attribute the processing of the emotional stimuli to the specific brain regions and frequency bands, we did not focus on these specific areas and analyzed every aspect of the EEG signal in all areas. We are interested in the inter-individual differences between subjects and we would like to test whether it is possible to find stable features in the EEG signal differentiating the emotional states for all subjects or whether we need to do the feature selection individually.

## II. Methods

Prior to the experimental procedure we administrated the EPQ-R personality test [6] to all participants to examine a correlation of its results with the EEG signal analysis. The experimental part of the research was based on the presentation of pictures eliciting different emotions. There were 100 pictures projected on the LCD screen in front of the subject. The experimental sample consisted of 23 subjects with the average age 25,6 years. The experiment time was 12 minutes for each subjects and it consisted of 2 minutes of baseline measurement and 10 minutes of the stimuli presentation.

There were four categories of the pictures, chosen from the IAPS database [7] and each category consisted of 25 stimuli. We selected pictures with the highest and lowest value of valence (pleasant or ugly) and arousal (boring or arousing). The mean average values for the categories were IAPS arousal = 2.37 for boring, IAPS valence =8.01 for pleasant, IAPS arousal = 7.14 for arousing and IAPS valence = 1.58 for ugly stimuli. The duration of each stimulus was 6 seconds and there was not delay between the picture presentation (because of the better results of ECG measurement). As the negative pictures should contamine the other categories we presented stimuli in fixed order. There are four categories following the presentation order: C1 – low arousal, C2 – high valence, C3 – high arousal, C4 – low valence. The set of pictures was presented to all subjects and we recorded their EEG activity.

The signal was recorded from 19 electrodes placed according to the 10-20 international system. The sampling rate was set to 250 Hz. After the recording we applied 50 Hz notch filter to the signal. Then we re-sampled it to 128 Hz and removed the isoline.

### *A. Feature Extraction*

The signal was adaptively segmented and the signal features were calculated. The process of adaptive segmentation is driven by amplitude- and frequency-dependent changes in

This work was supported by the research programme No. MSM 6840770012 " Transdisciplinary Research in the Field of Biomedical Engineering II". M. Macaš, M. Vavrečka, V. Gerla and L. Lhotská are with the Czech Technical University in Prague, Gerstner Laboratory, Technicka 2, Prague 6, Czech Republic (email: mmacas@seznam. cz)

EEG signal. The application of these parameters leads to division of the signal to the specific segments. This approach is presented in detail in [8].

As biological signals contain not only useful information, but also noise, it is necessary to transform the data to serve as an input for classification system. We implemented following algorithms to extract the features:

- statistical analysis (mean, standard deviation, skewness, kurtosis) of EEG and ECG
- power spectral density for classical band definitions in EEG (0.1-3 Hz, 3-7 Hz, 7-12 Hz, 12-20 Hz, 20-30 Hz, 30-40 Hz)
- statistical analysis derived from 1st and 2nd derivation of EEG
- statistical analysis (min, max, median, mean, std, skewness, kurtosis) applied to coefficients of wavelet transform (4 levels of decomposition; mother wavelet Daubechies 4)
- Shannon's entropy of wavelet transform details and approximations
- intra-hemispheric and inter-hemispheric EEG coherences between EEG electrodes
- correlations between electrodes
- heart rate variability (HRV)

We obtained 93 features for each EEG electrode, 8 features for ECG signal and we also calculated 127 intra hemispheric and inter-hemispheric correlation and coherence between electrodes. The total number of features is 1902 for each subject (19 EEG electrodes x 93 + 1 ECG electrode x 8 + 127 = 1902). All features were resampled to the constant time resolution (1 second). These feature data served as the input to the algorithms for feature selection and classification.

## B. Preprocessing

After the feature extraction, we obtained 6 data instances (each instance extracted from 1 second of the signals) for one stimulus (picture that should evoke a particular emotion). Totally, 25 stimuli from each of 4 classes were measured that gives 6x25x4=600 instances for one subject (person). Thus, data matrix 600x1902 was obtained, where each instance was classified into one of four classes (emotional states).

Before testing a classifier, we preprocessed the training data using the following procedures:

- **Removing outliers** - class by class a distance matrix is constructed and objects are removed that have a fraction 1/10 of their distances larger than the average distance in the class + 3 times the standard deviation of the within-class distances.
- **Normalization** - transform that shifted the data into origin (the mean value of each transformed is feature zero) and scaled the variances of each feature to 1.
- **Feature pre-selection** - the features were individually evaluated using inter/intra distance criterion and the best 100 features were further used. The inter/intra distance criterion [9] is distance-based class separability criterion, that is a monotonically increasing function of the distance between expectation vectors of different classes, and a monotonically decreasing function of the scattering around the expectations.
- **Feature selection** - subset of 7 features was selected using forward search algorithm and the inter/intra distance criterion.

## C. Classification

To provide the fair testing, we had to avoid the presence of two instances corresponding to same stimulus in training and testing set simultaneously. Each cross-validation fold was created using 5 randomly selected stimuli (5x6=30 instances) from each class that gave 4x30=120 instances at all. Thus, we have used a special case of 5-fold cross-validation.

The normalization and feature selection transforms were computed using training data. Thus, we obtained different feature subsets for different cross-validation folds. Such a method implies the fairness of classifier testing. The following procedure was followed for experimental testing

TABLE I
RESULTS OF INDIVIDUAL CLASSIFICATION

| Subject | E1 [%] | E2 [%] | E3 [%] | E4 [%] | E [%] |
|---|---|---|---|---|---|
| S1 | 5.3 | 34.0 | 35.3 | 24.7 | 24.8 |
| S2 | 34.0 | 27.3 | 41.3 | 49.3 | 38.0 |
| S3 | 16.7 | 2.0 | 5.3 | 10.0 | 8.5 |
| S4 | 22.7 | 38.0 | 49.3 | 26.7 | 34.1 |
| S5 | 4.7 | 9.3 | 22.7 | 10.7 | 11.8 |
| S6 | 2.0 | 44.7 | 67.3 | 32.0 | 36.5 |
| S7 | 24.0 | 38.7 | 32.0 | 17.3 | 28.0 |
| S8 | 15.3 | 14.0 | 15.3 | 24.7 | 17.3 |
| S9 | 25.3 | 47.3 | 31.3 | 32.6 | 34.2 |
| S10 | 6.7 | 20.0 | 28.7 | 36.7 | 23.0 |
| S11 | 8.7 | 42.7 | 35.3 | 6.0 | 23.1 |
| S12 | 22.0 | 10.0 | 22.7 | 12.7 | 16.9 |
| S13 | 12.7 | 5.3 | 10.0 | 20.7 | 12.2 |
| S14 | 36.0 | 53.3 | 44.7 | 10.0 | 36.0 |
| S15 | 6.7 | 32.0 | 33.3 | 33.3 | 26.3 |
| S16 | 14.0 | 8.7 | 34.0 | 13.3 | 17.5 |
| S17 | 17.3 | 18.0 | 16.7 | 16.0 | 17.0 |
| S18 | 12.7 | 24.0 | 46.7 | 15.3 | 24.7 |
| S19 | 38.0 | 47.3 | 24.7 | 29.3 | 34.8 |
| S20 | 10.7 | 22.7 | 24.0 | 22.0 | 19.8 |
| S21 | 10.7 | 17.3 | 6.7 | 8.7 | 10.83 |
| S22 | 22.7 | 33.3 | 26.7 | 12.0 | 23.7 |
| S23 | 11.3 | 36.0 | 50.7 | 36.0 | 33.5 |
| Mean | 16.5 | 27.2 | 30.6 | 21.7 | 24.0 |

Class errors for particular subjects. E1-4 denote the cross-validation error for classes 1-4 and E denotes the total cross-validation error.

TABLE II
AVERAGED CONFUSION MATRIX

| | | Estimated labels | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| True labels | 1 | 25.0 | 3.6 | 0.7 | 0.7 |
| | 2 | 3.5 | 21.8 | 3.3 | 1.4 |
| | 3 | 0.9 | 3.7 | 20.8 | 4.6 |
| | 4 | 0.6 | 0.9 | 5.0 | 23.5 |

Confusion matrix averaged over all subjects. The numbers denote average absolute occurrence in 30 instances.

of the classifier:

**FOR** each cross-validation fold **DO**

(1) **Remove outlayers** from training set
(2) **Normalize** the training set and apply the computed transform on the testing set
(3) **Pre-select** 100 features using individual selection and training set
(4) **Select** 7 of the 100 features using forward selection and training set. Select the 7 features from the testing set
(5) **Train** the classifier on the training set
(6) **Test** the classifier on the testing set

**END FOR**

(7) Compute the average cross-validation errors

After many preliminary experiments, we decided to use the Bayes classifier [10] and assume normal distributions of data from particular classes. The decision rule for such a classifier is: assign a feature vector $\vec{x}$ to class $\omega_i$ if $g_i(\vec{x}) > g_j(\vec{x})$ for all $j \neq i$, where

$$g_i(\vec{x}) = P(\omega_i \mid \vec{x}) = \frac{P(\vec{x} \mid \omega_i)P(\omega_i)}{P(\vec{x})}$$ is an aposterior probability.

For our case, we can assume identical prior probabilities $P(\omega_1) = P(\omega_2) = P(\omega_3) = P(\omega_4)$ and thus the decision rule can be reduced to maximization of $P(\vec{x} \mid \omega_i)$. The normal density has been assumed for its estimation. The decision rule leads to quadratic discriminant function.

## III. Results

The results of the classification testing are depicted in Table I. For each subject, cross-validation error (testing error averaged over the 5 cross-validation folds) was computed for particular classes (E1-4) and for the whole data. One can see that the error averaged over all subjects is 24% (see the last row of Table I), while for the best classified subject (S3) the error was 8.5%.

Furthermore, for each fold, a confusion matrix was obtained that shows how the instances from particular classes are confused. We have averaged the confusion matrices over all cross-validation folds and all subjects. The result is depicted in Table II.

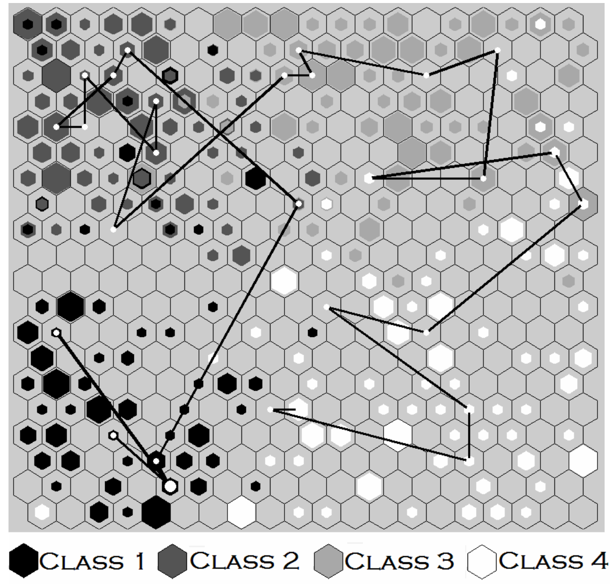


Fig. 1. The SOM visualization of the experiment with subject S3. The color hexagons represent the hits from each class and the black line is the trajectory over best-matching units computed for average of 30 adjacent instances.

Next, self-organizing map (SOM) [11] was trained on subject S3 data using 7 features selected on this subject in the first crossvalidation fold. The hit-plot is shown on Fig. 1. The hit-plot uses color hexagons whose sizes correspond to number of times a particular neuron is best-matching unit for a class. One can see that the SOM trained on subject S3 data shows that instances from a particular class are distributed in one area of the 7-dimensional input space. Moreover, one can see the trajectory (best matching units) in the Fig. 1 that shows how the person's brain activity changed over the time. The trajectory plot is created by averaging 30 adjacent instances and finding corresponding best matching units (white dots).

The list of features selected at least in one cross-validation fold for the subject S3 is depicted in Table III.

## IV. Discussion

We should imply the relation between the emotional category and classification error from the Tab. 2. The worst results for the category C3 (arousing pictures) should be attributed to the type of stimuli, because the most arousing pictures differ in the valence scale as they are ugly (e.g. mutilated bodies) or pleasant (e.g. erotic pictures). This ambivalence should negatively affect the classification. Better results should be accomplished by separate feature selection and classification on valence scale (category C2 and C4) and the arousal scale (C1 and C3). The confusion matrix in Tab. II confirms this hypothesis, because the category C3 (arousing pictures) has the highest confusion to the category C2 and C4 (pleasant and ugly).

As we were interested in finding regularities between emotional responses and the personality traits, we administrated the EPQ-R questionnaire to all participants prior to the experimental procedure. We compared the results of the classification with the results for the EPQ-R. The analysis uncovers the correlation between the classification error and the extroversion scale in the questionnaire. The subjects with the highest extroversion obtain the biggest classification error (approx. 30 percent). The lowest score in the extroversion scale resulted in the best classification ratio (approx. 10 percents). We can interpret these results in consistence with the psychological theories of emotion [12]. The introverts have a low threshold of excitation and tend to react to very small affective stimuli. It should result in biggest variance in the EEG activity and better classification. But this is just the preliminary hypothesis and has to be tested in the future research.

## V. Conclusion

We can conclude that the individual classification algorithms are able to detect emotional states with 75 percent accuracy. But there are still problems with the ability to generalize the results. It is necessary to calibrate this system separately for each participant with the known type of stimuli to identify the most salient features for the classification. The inter-individual variability documented in the previous studies [13] does not allow us to identify the same features for all subjects. It is also necessary to test the reliability of the calibration and its stability over longer time periods. These should be confirmed by testing the same subjects and applying the same methodology in the future.

It is also possible to eliminate the inter-subjective variability by the subtraction of the baseline (the normal EEG activity) from the signal. We did the baseline measurement, but did not apply the subtraction algorithm to the data, because it should not affect individual classification accuracy. The future development of this research will be focused on testing the reliability of this method and trying to extract the most frequent features able to discriminate between different emotional states.

TABLE III
Features selected for subject s3

| Feature | # |
|---|---|
| The mean second derivation of amplitude for appropriate segment in channel O2 | 5 |
| Maximum value of detail coefficients on the 5th level of decomposition (db4 wavelet) obtained from channel Cz | 5 |
| EEG coherence in beta band between T6 and O2 channels | 5 |
| EEG coherence in gama band between T6 and O2 channels | 5 |
| The mean second derivation of amplitude for appropriate segment in channel Fp2 | 4 |
| The mean value of the electrical resistance of the skin | 3 |
| Median of detail coefficients on the 5th level of decomposition (db4 wavelet) obtained from channel Cz | 2 |
| The minimum value of the electrical resistance of the skin | 2 |
| Maximum of detail coefficients on the 1st level of decomposition (db4 wavelet) obtained from channel F8 | 1 |
| The mean amplitude in channel Fz | 1 |
| Maximum of detail coefficients on the 5th level of decomposition (db4 wavelet) obtained from channel Cz | 1 |
| The mean of detail coefficients on the 5th level of decomposition (db4 wavelet) obtained from channel Cz | 1 |

All features selected at least once in the 5 cross-validation folds and the occurrence of each feature in the five folds.

## References


[1] R. J. Davidson and N. A. Fox , "Asymmetrical brain activity discriminates between positive and negative affective stimuli in human infants," *Science*, vol. 218, no. 4578, pp. 1235-1237, 1982.

[2] R. J. Davidson, P. Ekman, C. D. Saron, J. A. Senulis and W. V. Friesen, "Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology I," *Journal of Personality and Social Psychology*, vol. 58, no. 2, pp. 330-341, 1990.

[3] R. J. Davidson, "What does the prefrontal cortex do in affect: Perspectives in frontal EEG asymmetry research," *Biological Psychology*, vol. 67, no. 1-2, pp. 219-233, 2004.

[4] J. A. Coan and J. J. B. Allen, "Varieties of Emotional Experience during Voluntary Emotional Facial Expressions," *Annals of the New York Academy of Sciences*, vol. 1000, no. 1, pp. 375-379, 2003.

[5] J. A. Coan and J. J. B. Allen, "Frontal EEG asymmetry as a moderator and mediator of emotion," *Biological Psychology*, vol. 67, no. 1-2, pp. 7-49, 2004.

[6] H. J. Eysenck, S. B. G. Eysenck, *Manual of the Eysenck Personality Questionnaire*. London: Hodder and Stoughton, 1975.

[7] P. J. Lang, M. M. Bradley and B. N. Cuthbert, *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-8, University of Florida, Gainesville, FL., 2008.

[8] V. Krajča, S. Petránek, I. Patáková and A. Värri, „Automatic identification of significant graphoelements in multichannel EEG recordings by adaptive segmentation and fuzzy clustering," *Int. J. Biomed. Comput.*, vol. 28, no. 1-2, pp. 71-89, 1991.

[9] F. v. d. van der Heijden, R. Duin, D. d. de Ridder, and D. M. J. Tax, *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*. John Wiley & Sons, November 2004.

[10] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification (2nd Edition)*. Wiley, 2000.

[11] T. Kohonen, *Self-Organizing Maps*. Third, extended edition. Springer, 2001.

[12] C. J. Jung, *Psychologischen Typen.* Rascher Verlag, Zurich - translation H. G. Baynes, 1921.

[13] A. J. Tomarken, R. J. Davidson, R. E. Wheeler and R. C. Doss, "Individual differences in anterior brain asymmetry and fundamental dimensions of emotion," *Journal of Personality and Social Psychology*, vol. 62, no. 4, pp. 676-687, 1992.