Analysis of Event-Related Potentials for Emotion Recognition

Abstract. The primary objective of this study was to determine the feasibility of classifying emotions into three categories (positive, negative, and neutral) using event-related potentials (ERPs) for individual users. Visual stimuli from the International Affective Picture System (IAPS) database were utilized. Various features, such as signal samples, discrete wavelet transform, discrete Fourier transform, and discrete cosine transform, were computed from one-second electroencephalographic signal (EEG) segments following the presentation of the stimulus. For the classification task, a one-nearest neighbor classifier (1-NN) was employed. The research yielded a system for preprocessing and classifying emotions. The study involved eight participants. The experiments presented in this paper demonstrate the possibility of distinguishing emotions into three categories (pleasant, unpleasant, and neutral) for a single user, achieving an average accuracy level of 87%. However, when considering all users collectively, we achieved a classification accuracy of 96%.

Introduction

Recently, there has been a growing interest in the field of emotion recognition, which is a complex issue related to psychology. Emotions are commonly considered within the valence/arousal plane [1]. In this study, our focus was on utilizing event-related potentials (ERPs) of electroencephalographic signals (EEG) for emotion recognition. This method is considered highly reliable and capable of producing reproducible results [2]. The detection and recognition of emotions using EEG signals are actively developing fields of scientific research [3]–[6]. However, a significant portion of the research in this area is primarily theoretical and grounded in psychology [7]–[11]. Fewer works take a strict engineering approach, primarily focusing on signal processing and analysis methods. The study presented in [12] investigated ERPs with different valence values, employing a Morlet wavelet filter for feature extraction and support vector machine (SVM) for feature elimination. The paper described in [13] classified emotions such as happiness, surprise, fear, disgust, and neutrality using a combination of surface Laplacian filtering, wavelet transform (DWT), and linear classifiers. In [14], a novel architecture for discriminating emotions evoked by viewing pictures, utilizing biosignals from both the central and autonomic nervous systems, was proposed. In [2], it was discovered that the effect of emotion was sensitive to arousal in parietal electrodes and to both arousal and valence in frontocentral electrodes. Article [15] analyzed ERPs using spatiotemporal principal component analysis (PCA). In [16], a hybrid deep learning algorithm was proposed, which utilized convolutional neural network (CNN) layers for feature extraction on input data and combined them with long short-term memory (LSTM) networks for sequence prediction support. The CNN-LSTM classification with the ResNet152 model demonstrated high accuracy. In [17], deep learning analysis was employed, which overcame the challenges associated with hand-engineered feature extraction and selection.

The main objective of this study is to determine the feasibility of automatically classifying emotions into three categories (positive, negative, and neutral) for a single user using evoked potentials. Our aim is to develop a comprehensive system equipped with automatic artifact removal and effective algorithms for emotion classification. For visual stimuli, we utilized pictures sourced from the International Affective Picture System (IAPS) database [18]. During the experiments, we evaluated several preprocessing methods to enhance the quality of the EEG signal by removing artifacts. Features were computed from a one-second time windows of the EEG signal following the presentation of the stimulus. Signal samples were either used directly or transformed through methods such as discrete wavelet transform (DWT), discrete Fourier transform (DFT), or discrete cosine transform (DCT). Feature selection was performed using the t-test. In the classification stage, a 1-NN classifier was employed. The research study involved eight participants.

Materials and methods

a. Visual stimuli

One problem encountered during the research on detecting emotions in EEG signals was the creation of a representative database of visual stimuli (pictures) capable of eliciting the desired emotions. To address this, we meticulously selected images from the IAPS database, which offers a diverse range of pictures that affect the viewer to varying degrees. The validity of the database has been established through statistical surveys involving numerous individuals. In research on emotion detection in EEG signals, it is common to utilize extreme emotional stimuli. The pleasant pictures often depict sexual acts, while the unpleasant ones portray highly intense scenes such as injured accident victims. Our selection of images from the IAPS database covers a wide range of topics. The primary criterion for selection was the valence parameter for males. The images were categorized into three groups: pleasant, unpleasant, and neutral. Figure 1 illustrates the placement of the selected images on the valence/arousal plane. Table 1 presents the mean values and standard deviations of the valence and arousal parameters for each stimulus group.
The average luminance and their respective standard deviations were calculated for the images belonging to the three test classes to evaluate the influence of image brightness on classification accuracy. The mean luminance values for the neutral, unpleasant, and pleasant images were determined as 112.6 ± 42.8, 103.2 ± 34.7, and 113.2 ± 38.2, respectively (range 0 to 255).

### b. Visual stimuli

EEG signals were recorded from eight male participants, with an average age of 21 years, all of whom were students at the Warsaw University of Technology. None of the participants had a history of neurological diseases. During the experiment, the participants were instructed to observe the presented stimuli, which consisted of sequentially displayed pictures. A fixation cross was shown on the screen before each image. The pictures were categorized into three sets: pleasant, unpleasant, and neutral. The presentation of the pictures occurred in a random order, as depicted in Figure 2.


### c. Feature extraction and classification

An integral aspect of analyzing event-related potentials (ERPs) involves establishing the baseline of the EEG signal. In this particular scenario, the baseline was determined by subtracting the average value of the half-second segment of the EEG signal preceding the stimulus onset from the ERP signal. The selection of the half-second time period was set experimentally and resulted in favorable classification accuracy. The next significant question that arose was the duration of the signal to be analyzed after stimulus presentation. Based on numerous experiments and studies in the literature, we determined to analyze a one-second-time interval following the stimulus arrival.

Multiple feature extraction methods were investigated in the study. One simple and intuitive approach involved considering the shape of the EEG signal in the time domain, where the signal samples themselves were regarded as features. The sampling rate was set at 256 S/s, resulting in 256 possible features within a second of an ERP signal. Additionally, several other feature extraction methods were examined, including:

- The approximation of the discrete wavelet transform (DWT) on the first level of composition.
- Absolute values of discrete Fourier transform (DFT) coefficients.

Each of these methods possesses specific parameters and properties, which are described in Table II. The features were calculated from a one-second time window of the EEG signal following the stimulus arrival.

### Table 2. Feature extraction methods

<table>
<thead>
<tr>
<th>Feature extraction method</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samples</td>
<td>Different time intervals</td>
<td>Samples from the range: 0-64, 64-128, 128-192, 192-256</td>
</tr>
<tr>
<td>DWT</td>
<td>Wavelet type</td>
<td>db2, db4, db5, db7, sym2, sym4 and other</td>
</tr>
<tr>
<td>DFT</td>
<td>DFT size</td>
<td>256, 128, 64</td>
</tr>
<tr>
<td>DCT</td>
<td>Coefficients and absolute values of the coefficients</td>
<td>Coefficients and absolute values of the coefficients</td>
</tr>
</tbody>
</table>

We evaluated multiple ranking methods for feature selection, and the t-test proved to be the most effective. For each feature extraction method, the features were individually selected. The experiments demonstrated that utilizing the t-test algorithm with 30 features produced satisfactory results. Since the ranking methods operate as binary classifiers, selection can only be conducted between pairs of categories. Considering our three classes, a total of 120 features were chosen. Subsequently, we narrowed down this selection to the top 40 unique features (as several were repeated for class pairs). These 40 features were employed for classification in subsequent experiments.

### Results and discussion

We considered multiple methods for learning and testing classifiers. One of them was the cross-validation test. However, for a 10 cross-validation test, the number of signals to average (ERPs, 1-sec signal segments) would be too small for classification. At the same time, utilizing single representations of EEG signals (without averaging) for feature extraction proves disadvantageous due to the excessively low signal-to-noise ratio. The only effective method we tested involved randomly selecting EEG signals.
for training and learning sets. Initially, all the data (a collection of 1-sec signal sections - ERPs) was randomly divided into two separate subsets for training and testing. Subsequently, the data from both subsets were averaged. For the classification task, a 1-NN classifier was employed. The similarity between testing and training features was measured using the Euclidean distance. The learning data was drawn from the first (averaged) subset, while the testing data originated from the second subset. This process was repeated multiple times to create training and testing examples. The classification accuracy was then averaged over multiple algorithm runs. This approach not only facilitated the elimination of differences in EEG signals that occur between sessions for a single user but also addressed variations stemming from user movements, changes in skin-electrode contact conductivity, or habituation to stimuli. By implementing a random selection of training and testing data followed by result averaging, we developed an effective classifier learning method.

The experiments were conducted for each of the eight participants. Figure 3 presents the classification results for each participant across the three categories (pleasant, unpleasant, neutral) using the Sample feature extraction method, which yielded the best classification results. The accuracy ranges from 0.85 to 0.95, varying among the participants.

We also conducted experiments to determine the most valuable electrodes for the classification process. Through numerous tests, we identified the seven electrodes that yielded the best results: O2, P4, P3, Pz, CPZ, Oz, and O1. The outcomes obtained using these seven electrodes were comparable to those achieved with all 16 electrodes. The obtained classification accuracy results are very difficult to compare with other studies presented in the literature. This is due to the fact that each experiment was conducted under different conditions, recording different EEG signals, and stimulating the user with completely different stimuli. In [19], the performance of the presented method is evaluated by classifying emotional valence into three levels: extremely negative, moderately negative, and neutral, using support vector machine. The highest accuracy achieved in the three-class classification is 77.5%. In [20], the paper focuses on classifying emotions into four classes on a valence/arousal plane. The average event-related potential (ERP) attributes and differentials of average ERPs obtained from the frontal region of 24 individuals were utilized for the emotion classification. The results of the subject-independent four-class emotion classification ranged from 67% to 83%. By employing three classifiers, a mid-range accuracy of 85% was achieved.

**Conclusion**

The experiments described in the literature typically involve averaged EEG signals from a large number of users. This approach allows for the identification of general psychological patterns across many individuals. However, it does not address whether it is possible to detect emotions for a single user. The experiments presented in this paper demonstrate that it is indeed possible to distinguish emotions into three categories (pleasant, unpleasant, and neutral) for a single user, achieving an accuracy level of 87%. The best classification results can be achieved using the Samples feature extraction method (accuracy 95%) and the DWT feature extraction method (accuracy 94%). The methods employed, including incorporating the signals from all users (around 320 for training and 320 for testing the classifier), we were able to achieve a classification accuracy of 96%. The increased number of stimuli led to more efficient classification.
preprocessing, feature extraction, and selection, are applicable universally and can be implemented in an automatic system with minimal operator involvement. Significantly improved results (attributable to increased averaging) can be obtained by utilizing cumulative data from all users.

Authors: dr hab. inż. Andrzej Majkowski, Warsaw University of Technology, Institute of Theory of Electrical Engineering, Measurements and Information Systems, Koszykowa 75, 00-662 Warsaw, E-mail: andrej.majkowski@pw.edu.pl; dr hab. inż. Marcin Łołodziej, Warsaw University of Technology, Institute of Theory of Electrical Engineering, Measurements and Information Systems, Koszykowa 75, 00-662 Warsaw, E-mail: marcin.lolodziej@pw.edu.pl; prof. dr hab. inż. Remigiusz Jan Rak, Warsaw University of Technology, Institute of Theory of Electrical Engineering, Measurements and Information Systems, Koszykowa 75, 00-662 Warsaw, E-mail: remigiusz.raki@pw.edu.pl;

REFERENCES


