P300 Wave Detection Based on Rough Sets

Sheela Ramanna[1] and Reza Fazel-Rezai[2]

[1] Department of Applied Computer Science, University of Winnipeg,
Winnipeg, Manitoba R3B 2E9 Canada
s.ramanna@uwinnipeg.ca

[2] Department of Electrical and Computer Engineering, University of Manitoba,
Winnipeg, Manitoba R3T 5V6 Canada
fazel@ee.umanitoba.ca

Abstract. The goal of P300 wave detection is to extract relevant features from the huge number of electrical signals and to detect the P300 component accurately. This paper introduces a modified approach to P300 wave detection combined with an application of rough set methods and non-rough set based methods to classify P300 signals. The modifications include an averaging method using Mexican hat wavelet coefficients to extract features of signals. The data set has been expanded to include signals from six words and a total of 3960 objects. Experiments with a variety of classifiers were performed. The signal data analysis includes comparisons of error rates, true positives and false negatives performed using a paired t-test. It has been found that the false negatives are better indicators of efficacy of the feature extraction method rather than error rate due to the nature of the signal data. The contribution of this paper is an in-depth study P300 wave detection using a modified averaging method for feature extraction together with rough set-based classification on an expanded data set.

Keywords: Brain computer interface, EEG signal classification, Mexican hat wavelet, P300 wave detection, feature extraction, rough sets.

1 Introduction

Brain Computer Interface (BCI) involves monitoring conscious brain electrical activity, via electroencephalogram, (EEG) signals, and detecting characteristics of brain signal patterns, via digital signal processing algorithms, that the user generates in order to communicate with the outside world. BCI technology provides a direct interface between the brain and a computer for people with severe movement impairments. The goal of BCI is to liberate these individuals and to enable them to perform many activities of daily living thus improving their quality of life and allowing them more independence to play a more productive role in society and to reduce social costs. Considerable research has been done on BCI (see, e.g., [4,7,8,12,20]). One of the benefits of the P300-based BCI is that it does not require intensive user training, as P300 is one of the brain's "built-in" functions. A particular challenge in BCI is to extract the relevant signal from the huge number of electrical signals that the human brain produces each second. In

J.F. Peters and A. Skowron (Eds.): Transactions on Rough Sets V, LNCS 4100, pp. 207–223, 2006.

addition, it is also critical for any BCI system to be of practical use, the number of channels used to extract the signals be kept to a minimum. Event related potentials (ERPs) are psychophysiological correlates of neurocognitive functioning that reflect the responses of the brain to changes (events) in the external or internal environment of the organism. ERPs have wide usage for clinical-diagnostic and research purposes. In addition, they have also been used in brain computer interfaces [5]. P300 is the most important and the most studied component of the ERP [4]. Previous work in classifying signal data with rough set methods has shown considerable promise [9,19]. Also, more recently a gradient boosting method for P300 detection has been used by Hoffmann *et al.* [16]. The differences between our two approaches are significant. First, Hoffmann *et al.* have designed their own experiment to collect signal data from 10 channels and to detect P300 on BCI data. Second, gradient boosting was used to stepwise maximize the Bernoulli log-likelihood of a logistic regression model. Ordinary least squares regression was used as weak learner with a classification accuracy between 90-100% for the BCI data.

In this paper, a slightly modified method for P300 wave detection is introduced. The efficacy of P300 detection consists of two components: feature extraction and classification. For feature extraction, Mexican hat wavelet coefficients provide good features when averaged over different scales [8]. We have used a slightly different averaging method based on Mexican hat wavelet coefficients. The channels used to extract the signals are also changed. The range of methods used for the classification of features values now include three rough-set based methods and three non-rough set based methods. In addition, we have used a more extensive data set as compared to our previous work [9]. The experiments were conducted on the data set provided by the BCI group at the Wadsworth Center, Albany, NY. This data set represents a complete record of P300 evoked potentials recorded with BCI2000 using a paradigm described by Donchin et al., 2000. Since the expected response to a particular character (and subsequently the word) is already known, supervised learning methods are ideal for predicting the correct character sequence. In our previous work [9] both standard supervised and a form of sequential character-by-character classification was used. More recently, layered or hierarchical learning for complex concepts has been successfully applied to data from road-traffic simulator [22] and classification of sunspot data [23]. In both these cases, complex concepts are decomposed into simpler but related sub-concepts. Learning at a higher-level is affected by learning at a lower level where sub-concepts are learned independently. Incremental learning is another form of learning [33] where the structure of the decision table is changed (updated) incrementally as new data are added to the table rather than regenerating the whole table. In other words, learning occurs on a hierarchy of decision tables rather than on a single table. Our work differs from both of these methods since (i) the concept (word recognition) is simple, and (ii) our table does not change over time. Upon further experimentation, we have found that there is no gain in using the character-by-character classification approach. It should also be noted the number of features and the number of channels used