Automatic Classification of Visual Evoked Potentials Based on Wavelet Analysis and Support Vector Machine

M. Akbari 1, R. Azmi 2

1Islamic Azad University, Shahr-e-Qods Branch, m.akbari@qodsiau.ac.ir
2Faculty Member, Alzahra University, azmi@alzahra.ac.ir

Abstract—In this paper 54 normal visual evoked potential signals and 16 abnormal VEP signals were used. These signals are presented in a 512 element vector. By applying the wavelet, after compression and using the threshold level, the number of features was reduced. Afterwards these features were classified with the k-Nearest Neighbor (KNN) and Support Vector Machine (SVM) methods and the precision of these two methods were compared. The results of this test indicate the best error rate in k-Nearest Neighbor method 23.07% while the error rate in support vector machine method is 30.76%. To reduce the error rate in these methods, we considered 100 features extracted in the Daubechies method. A suitable method of feature selection indicate that with choosing the proper features the error rate in k-Nearest Neighbor and Support Vector Machine methods reduce to 5% and 20% in sequence. Furthermore with applying Leaving one out method on 8 features the final error rate was evaluated 6.25%.

Keywords—Visual Evoked Potentials, Wavelet transform, Support Vector Machine Classifier, k-Nearest Neighbor Classifier, Biosignal processing

I. INTRODUCTION

Visual Evoked Potential (VEP) is a signal providing information produced by ganglion cells from the visual system. This information is transmitted by the visual nerves and the reflex is recorded in the visual cortex. In comparison to electroretinogram and EOG which are an electro tonic conduction and vary with different light intensities, visual evoked potentials are kinds of action potential which its production imitates from the all or none principle [1]. In a study accomplished by Palaniappan entitled “Discrimination of Alcoholic Subjects using Second Order Autoregressive Modeling of Brain Signals Evoked during Visual Stimulus Precision” it has been specified that with using neural network simplified fuzzy ARTMAP (SFA) methods, neural network Multilayer Perceptron backpropagation (MLP-BR) method and linear diacritic (LD), an error equivalent to 11.9%, 2.8% and 2.6% in sequence for each method is presented [2].

In this research this fact is presented that by applying some improvements on the classification of alcoholics and non-alcoholics and increasing the data collection to 3560 visual evoked potentials from 102 cases, the performance will increase from 94.49% to 98.71%. The k-nearest neighbor method was not the only method used in this method; previous to it the pass-band width of the used filter was increased and in continuance the Multiple Signal Classification algorithm was used in order to obtain the power of the dominant frequency in gamma band VEP signals. In the last stage the k-Nearest Neighbor in which k was equivalent to 1,2,3,4 and 5 was utilized.

In a research entitled Multi-channel noise reduced visual evoked potential analysis it is specified that if Phase zero of butterworth digital filter in used in order to extract gamma-band power in a 30 to 50 Hz spectra range from the visual evoked potentials with reduced noise, the fuzzy ARTMAP (FA) method with the aid of noise reduction application (PCA) will have an accuracy equivalent to 92.50%, whereas without the aid of this application it will have an 83.33% precision [3]. The Wavelet Transform is an analysis tool which is used in many cases related to signal analysis and recognition. Wavelet is a small wave which its energy is centralized in a small area and is a convenient tool for transitory phenomenon. Wavelet has a minimum oscillation which descends to zero and this descending is bounded in positive and negative directions. This characteristic enables the wavelet to be supple and behave like a function [4].

Clinical examinations have shown that the multi-resolution wavelet method can estimate and identify the delay peak for visual evoked potentials. This task requires fewer tests from the total average. The most benefits of the wavelet are that this transform is able to reach a convenient approximation of time discontinuities with a quick pace. Furthermore it can also appoint basic elements in both time and frequency domains. With the assistance of these benefits, the wavelet analysis can detect abnormalities and irregularities. One of the important characteristics of the wavelet transform is that its output has a small frequency and time boundary, which has focused the attention of many researchers in visual evoked potentials medical field [5].

For the purpose of analyzing biosignals such as brain signals one of the best methods is benefiting from the wavelet analysis in order to omit artifacts and extract data. This transform is convenient when revealing transitory events such as epilepsy and has a good resolution in time and frequency.

In image processing also, the wavelet is used in applications such as image compression and recognition, extracting certain
important areas from an image with a higher resolution, noise elimination and image recording [6].

Another part of the suggested algorithm includes an appropriate method for classification, hitherto many methods are used for data classification. Among these methods we can mention methods based on neural networks, k-nearest neighbor (KNN), support vector machine (SVM), and RBF (Radial Based Function) networks. In this research we have used k-nearest neighbor and support vector machine classifiers.

Although the support vector machine method is harder to implement in comparison to linear classifiers and neural network and uses a longer performance time, in general because of its higher precision in classification it has drawn a lot of attention. For example in a study which was carried out on the classification of newborns electroencephalography indicates better results in performance than neural networks [7].

In another study which was carried out with the purpose of classifying single-trial potentials evoked by imitating natural reading with v-SVM method, the results indicate that the precision of classification were 92.6%, 88.2% and 93.5%. These values are a consequence of quantitative dispersal in the results. The values are the result of 10 times testing. All of the experiments in this study are divided into two equal sets: the training set and the test set; each of the mentioned sets include testing 50% answers of the target samples and 50% answers of the non target samples [8].

The k-nearest neighbor is also among the methods used for signal processing such as VEP signal processing. In a research carried out by Palaniappan entitled Improved Automated Classification of Alcoholics and Non-alcoholics, in the first stage a larger set of data consisting of 3560 visual evoked potential signals from 62 alcoholics and 40 non-alcoholics were used. In the next stage the pass-band width of the used filter was increased and in continuance the multiple signal classification algorithm was used in order to obtain the power of the dominant frequency in gamma band VEP signals. In this study the k-nearest neighbor method was used in order to locate the nearest neighbor. In this paper k was varied from 1 to 5. This method reduces the design complexity. The maximum performance for the k-nearest neighbor method was 98.71% hitherto with using the basic method discussed previously this value was 95.73% [9].

II. METHODOLOGY

1. The Wavelet Method

In this study visual evoked potentials feature were extracted with using Haar and Daubechies wavelet. Because of the large number of features there was the need to provide a method which enables us to bring near the features with less effect on classification, to zero and consider only to the features with a higher effect as a distinction factor. Therefore after applying the wavelet with the aid of compression and with choosing an appropriate threshold the features were reduced to 100.

Another application of the wavelet transform is lossy compression. This compression consists of the following three basic stages [10]:

- A wavelet bank filter converts the image to wavelet coefficients.
- The obtained coefficients are quantized.
- Finally the entropy compressor is applied and the image is compressed

Choosing a filter and the attributes of the corresponding wavelet is dependent on the compression algorithm (quantization method and assigning the bits)

2. K-Nearest Neighbor Method

After applying the wavelet method we classify each signal in one of the two groups of normal and abnormal cases with the aid of k-nearest neighbor method. K-Nearest Neighbor method is a method used for classifying objects based on training samples nearer to the features’ space. This method is a kind of learning technique based on samples which only the completely local function is calculated and all calculations are delayed till classification is brought to an end. An object is relegated to a class which is standard among its neighbors, while it is classified on the bases of the majority of its neighbors. In cases which our goal is to classify the samples in two classes it is better to choose an odd number for k to prevent continuous votes [10].

Considering this fact that the results of classification were acknowledged before testing, we were able to calculate this methods error rate in differentiating the normal and abnormal signals from each other correctly. In this method different values were used for k. These values were 1, 3, 5 and 7 which the best results belonged to k=1. The results of applying this method are presented in table I.

<table>
<thead>
<tr>
<th>Wavelet Basis</th>
<th>KNN method</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K=1</td>
<td>K=3</td>
</tr>
<tr>
<td>Haar group 1</td>
<td>32.7</td>
<td>42.3</td>
</tr>
<tr>
<td>Haar group 2</td>
<td>36.5</td>
<td>40.3</td>
</tr>
<tr>
<td>db2, d=4, σ=100</td>
<td>21.5</td>
<td>30.7</td>
</tr>
<tr>
<td>db2, d=4, σ=100</td>
<td>36.7</td>
<td>53.8</td>
</tr>
<tr>
<td>db2, d=2, σ=133</td>
<td>30.7</td>
<td>28.8</td>
</tr>
</tbody>
</table>

3. Support Vector Machine Method

In this research in addition to the k-nearest neighbor method, the support vector machine method was also reviewed. In this method the chosen coefficients of the wavelet, and in other

M. Akbari, Reza Azmi
words, the features were addresses as the inputs of this network, among which some of the samples were randomly chosen as the training group. Applying this method the samples were classified in two groups: normal and abnormal. The error rate of this method is calculated based on the cases which were set in an incorrect group instead of their own group. (Table 1)

If we have a training set with double sample labels \((x_i, y_i)\) while \(i=1, \ldots, l\) in which \(x_i \in \mathbb{R}^d\) and \(y_i \in \{1, -1\}\), the support vector machine requires a solution for the following optimizing problem:

\[
\min_{w, b, \xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \\
\text{subject to } g_i(w^T \phi(x_i) + b) \geq 1 - \xi_i,
\]

\[
\xi_i \geq 0.
\]

Hither the training vectors \(x_i\) is mapped to an upper dimensional space. Then the support vector machine finds a linear discriminator with a maximum margin in this upper dimensional space. \(C>0\) is the required parameter for regarding the error effect. Furthermore \(K(x_i, x_j) = \phi(x_i)^T \phi(x_j)\) is the kernel function. The kernel function is normally of the following functions:

- Linear: \(K(x_i, x_j) = x_i^T x_j\)
- Polynomial: \(K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0\)
- Radial Basis Function (RBF): \(K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0\)
- Sigmoid: \(K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)\)

4. Leaving- One- Out method

This method is a straightforward technique for estimating error rate in classifiers. Because of its high cost in calculation it is often used in cases with a small number of samples, while wanting to achieve a high precision in evaluation. As we had a relatively fewer samples of abnormal cases, this test was used in order to register more precise results. In this method if we have \(n\) samples, the classifier generates results based on \(n-1\) samples, the rest of the samples. This task is repeated \(n\) times and each time the classifier is designed with leaving one of the samples out. In other words each sample is chosen as the sample which must be tested and in each stage nearly all samples are used for designing the classifier. The error rate is the number of errors on the samples used in the single-trial divided by \(n\). There are a lot of testimonials which indicate the advantage of this method. If there are a large number of samples in this method it is probable to encounter a high calculation cost [12].

III. THE EXAMINATION OF THE SUGGESTED ALGORITHM

1. Test Phases and Results

In this survey 70 visual evoked potential signals were used to examine the used classification methods’ error. Among these signals, 54 signals were from 54 normal cases and 16 abnormal cases. Each signal formed a 512 element vector. The results of applying each method after extracting the features with the aid of different wavelet methods are introduced in table 1. In figure (1) a sample of the waveform of visual evoked potentials signals are presented.

2. Examination and Analysis of Results

Considering the results presented in table 1 we understand that with changing the feature extraction method from Haar, the simplest method, to Daubichies the error has reduced. Furthermore with reducing the size of the coefficients matrix in db2 method with decomposition level 4 in all kinds of k-nearest neighbor with different values of \(k\) and support vector machine method the error rate has increased. The best results are generated from the db4 with decomposition level 2 method.

In addition to the results presented above we comprehend that the k-nearest neighbor method with \(k\) equal to 1 indicates less error in comparison to this method with 3, 5 and 7 neighbors and it has also shown better results in comparison to support vector machine method. The error rate in this technique is 23.07\%, while having an error equivalent to 30.76\% in support vector machine method.

3. The Effect of Feature Extraction Process

In order to survey the effect of appropriate feature selection among the wavelet coefficients (after compression and eliminating based on threshold level and choosing the prime 100 features) in the improvement of the final results we used a tentative algorithm in feature selection [13].

First, based on this algorithm we chose the best feature and in each stage we added a new feature so that the final structure will have a better result in recognition. This process is continued till we have a proper number of features. Applying this method to the features obtained in the Duabechies method resulted in 8 chosen features for the kNN (\(k=1\)) method. Figure 3 introduces the results of applying this technique. The error rate in this method for the KNN and Support Vector Machine methods was improved to 95\% and 80\% in sequence.
4. Final test, based on Leaving- One- Out method
As specified before with a small number of samples in one class it is an advantage to use the Leaving- One- Out method in order to obtain a better precision in evaluating the algorithm. This method is applied with the features selected in the last mentioned process and results show that error in this method with surveying the abnormal cases is 6.25% which indicates that the suggested algorithm has 93.75% precision.

![Figure 2. results’ graph based on the number of feature in the k-nearest neighbor method](image)

V. CONCLUSION

The research carried out indicates that the features resulted from the wavelet on visual evoked potentials (VEP) if chosen properly can be used for distinguishing the normal and abnormal cases from each other. According to the tests data, the 8 features selected from the db4 wavelet has the best precision in this approach. The estimation in this state with applying Leaving- One- Out is about 94%. In continuation of this study we will try to examine the results of recognition with the effect of noise in signals as well as testing other classifiers.

REFERENCES

[10] Miah M., Improved k-NN Algorithm for Text Classification., Department of Computer Science and Engineering University of Texas at Arlington, TX, USA
[12] Weiss Sholom M. and Kapoulas Ioannis., An Empirical Comparison of Pattern Recognition, Neural Nets, and Machine Learning Classification Methods, Department of Computer Science, Rutgers University, New Brunswick, NJ 08903